Volume 7 APRIL, 1936 Number 4

PHYSICS

A Journal of General and Applied Physics

# The Dielectric Constant of Air at Radiofrequencies

L. G. HECTOR AND H. L. SCHULTZ, *Department of Physics, University of Buffalo*

(Received November 16, 1935)

The heterodyne beat frequency method of determining dielectric constants has been applied to the measurement of the dielectric constant of air with an applied frequency of field of 900,000 cycles per second. Radio broadcast signals, stripped of modulation, were used as sources of alternating current of constant high frequency. An electron coupled oscillator was used in the circuit containing the test condenser. Both sources of high frequency were isolated from the detector through buffer stages of amplification. Determination of the beat frequencies was made by photographic recording with a two string oscillograph. A dielectric constant of $1.00058986 \pm 0.00000050$ is found for air at normal pressure and temperature. This result is in close agreement with the results of other investigators for lower frequencies of applied field and for static measurements.

## INTRODUCTION

IN the course of preliminary work on the effects of intense electric fields on the dielectric constants of gases, we have applied the heterodyne beat frequency method to the measurement of the dielectric constant of air when the frequency of alternation of the applied field is 900,000 cycles per second.

Basically this method of measuring dielectric constants is old. It consists in mixing a current of high but unvarying frequency in a detector with current supplied from a variable frequency oscillator. The detector then delivers a current with the difference or beat frequency to some type of frequency indicating instrument.

The variable frequency oscillator circuit includes a condenser whose capacity may be changed by the withdrawal or insertion of the material whose dielectric constant is to be determined. This capacity change may be compensated for by altering the setting of another condenser in the circuit; or the changes in capacity may be computed from the change in frequency of the beat note.

The principal experimental difficulties with the method are:

1. Maintaining the desired constancy of frequency of a standard high frequency source.
2. Maintaining the constancy of a variable frequency oscillator with respect to causes other than the change in the capacity of the test condenser.
3. Avoiding the tendency of the two sources to synchronize when the beat frequency approaches zero.

A solution to the first of these difficulties is the use of a crystal controlled oscillator with temperature control for the crystal and this type of source for a standard high frequency current is now in general use by experimenters in this field.

The third difficulty, that of synchronization, may be handled by radiation coupling to the detector from the two sources of high frequency current as was reported by Bryan and Sanders[1] in 1928.

It can also be solved by the use of buffer stages of amplification such as have now come to be accepted technique in high frequency radio engineering.

Solution to the second difficulty, that of the constancy of the variable frequency oscillator to causes other than changes in the test condenser, is now possible to a greater extent than formerly

[1] A. B. Bryan and I. C. Sanders, Phys. Rev. **32**, 302 (1928).

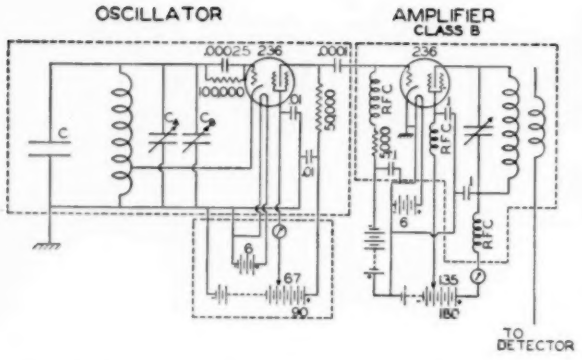


FIG. 1. Oscillator and amplifier class B.

through the development of multigrid tubes, special oscillating circuits for use with such tubes, and buffer stages to isolate the oscillator from the effects of any possible load variations in the output circuit.

## EXPERIMENTAL METHOD

We have attempted to incorporate the improvements outlined above in this well-recognized method for the measurement of dielectric constants with the hope that sufficiently improved performance might be obtained to enable us to look for dielectric saturation phenomena in gases. Largely as an attempt to see what precision could be attained we have made a series of measurements on the dielectric constant of air and wish to report the results of these measurements as well as the details of the method in the present paper.

Fig. 1 shows the circuit diagram of the variable frequency oscillator and its buffer stage. The oscillator is the so-called electron coupled type. Essentially it is a Hartley oscillator where the grid of the tube used ordinarily as a screen is used as the plate. Output power, however, is taken from the actual plate of the tube. An analysis of this type of oscillator and the reasons to which its constancy of frequency may be attributed have been given at length in papers by F. B. Llewellyn[2] and by J. B. Dow.[3]

The oscillator is resistance-capacity coupled to a buffer stage which in turn supplies power to the detector circuit. The buffer stage has its grid biased for class B operation and the output circuit may be tuned either to the fundamental or to an harmonic of the oscillator. In the work reported in this paper the second harmonic was used.

The entire oscillator unit was built into a shield box made of 3/8 inch sheet aluminum. Such heavy construction was found necessary to insure the required mechanical and hence electrical capacitative stability. The buffer stage was built into a separate aluminum shield and the two were kept in a large wood box to improve temperature control.

Filament and plate batteries were shielded in a copper lined box.

In Fig. 1 the test condenser is $C$. $C_b$ is a vernier condenser used to set the oscillator at approximately the frequency desired. $C_a$ is a smaller vernier condenser used for fine adjustment of the oscillator to zero beat with the standard high frequency.

The test condenser was of the multiple plate parallel plate type. The plates were made of brass and the whole unit was housed in a container turned from solid brass.

The source of standard high frequency was the broadcast signal from a radio station (WBEN) equipped with late type temperature controlled crystal oscillators. A method for removing the modulation from such signals so that the presence of program need not interrupt the use of the carrier for laboratory measurements is described in another paper by the authors.[4]

The two high frequencies were mixed in the detector amplifier circuit shown in Fig. 2. The output of this circuit feeds a standard string oscillograph.

Some idea of the constancy both of this standard signal and of the local oscillator may be gleaned from frequency data of which the following is typical. The beat note between the second harmonics of the standard frequency (900,000) and that of the local oscillator was set at about 100 cycles and the actual beat frequency accurately determined at 20-minute intervals. The values found were 99.94, 99.69, 99.00 showing a variation of about 1 cycle per second in 1,800,000 cycles per second in 40 minutes.

[2] F. B. Llewellyn, Proc. I. R. E. **19**, 2063 (1931).
[3] J. B. Dow, Proc. I. R. E. **19**, 2095 (1931).
[4] L. G. Hector and H. L. Schultz, Rev. Sci. Inst. **7**, 139 (1936).

## Experimental Procedure

Before taking any data the oscillator was allowed to run for a period of ten to twelve hours. The test condenser was then evacuated and the local oscillator set so that its second harmonic (1,800,000) gave zero beat with that of the second harmonic of the standard signal. This condition could be observed by the motion of the needle of a d.c. milliammeter in the plate circuit of the amplifier connected to the detector (see Fig. 2).

Air was now allowed to enter the test condenser after passing through a solution of KOH for removing $CO_2$ and through two large reservoirs packed with $CaCl_2$ for removing water vapor. For the change from vacuum to atmospheric pressure, the change in the frequency of the oscillator was approximately 215 cycles per second. This corresponds to a frequency of about 430 cycles per second between the second harmonic frequencies.

The actual frequency was determined precisely by means of photographic recording with a two string oscillograph where the beat note was recorded by one galvanometer and a timing wave by the other. The timing employed was that of the 60-cycle electric power supply.

Checks of the frequency deviations of this source from 60 cycles per second show that the variations seldom exceed 0.05 cycle per second and that they are as often above as below the assigned value.

## Determination of Stray and Effective Capacities in the Circuit

In computing the change in capacity of the test condenser from beat frequency data it becomes necessary to know the ratio of the stray capacity in the oscillatory circuit to that of the effective capacity of the test condenser. This necessity arises from the fact that obviously the stray capacity of the circuit does not change when gas is introduced into the test condenser.

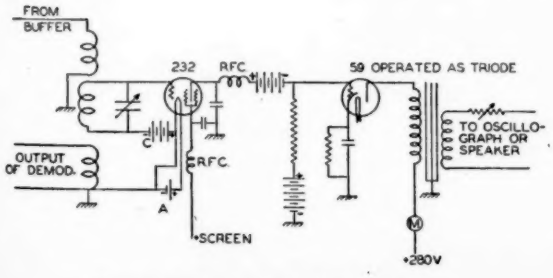


Fig. 2. Detector and A.F. amplifier.

The capacity of the test condenser was first determined by means of substituting a new, calibrated, type 222F General Radio air condenser in the circuit. The mean of several determinations gave a value of 518.4 mmfd. It was estimated that as much as 1.5 mmfd. of this capacity was so located as not to be effected by the introduction of gas. Hence the effective value of the test capacity is 516.9 mmfd.

For determining the ratio of stray to effective capacity for the entire circuit the following procedure was employed.

As in the experimental procedure described above, the frequency change of the variable frequency oscillator, originally at 900,000 cycles per second, was observed for an air pressure change in the test condenser from vacuum to a value near atmospheric.

Capacity from a variable standard condenser was now added in parallel to the test condenser until the frequency with the test condenser evacuated was 700,000 cycles per second. (Standard frequency was obtained from the signal of radio station WLW.) The frequency change on filling the test condenser with air to the same pressure as before was then determined.

From a knowledge of the two original oscillator frequencies, the two beat frequencies, the effective value of the test condenser, and the amount of added capacity for the second series of data the desired ratio between the stray capacity of the circuit and effective capacity of the test condenser, $R$, can be determined. The value for our circuit was found to be 0.09189.

## Calculation of the Dielectric Constant

If we write $\epsilon$ for the dielectric constant of the gas under the conditions of pressure and temperature at which the measurement is made, we can compute the value of $(\epsilon-1)$ from a knowledge of the ratio, $R$, and the change in frequency of the oscillating circuit when the surroundings of the test condenser are changed from vacuum to this pressure of gas.

Equations showing this relation may be developed as follows. The natural frequency, $f$, of the oscillating circuit is given by

$$f=1/2\pi(L(C+c))^{\frac{1}{2}}, \quad (1)$$

where $L$ is the inductance, $C$ the effective value of the test condenser, and $c$ the stray capacity of the circuit.

$$df/dC=-1/4\pi L^{\frac{1}{2}}(C+c)^{\frac{3}{2}}.$$

Remembering that $dC=C(\epsilon-1)$ and that $R=c/C$, we may write

$$\epsilon-1=-2df(1+R)/f. \quad (2)$$

Since the change in frequency is rather large to be treated in the manner of a differential, a more precise expression for $(\epsilon-1)$ may be obtained by expanding (1) by the method of increments. The result obtained in this manner is

$$\epsilon-1=(1+R)[1/(1+\Delta f/f)^2-1]. \quad (3)$$

Results published in this paper have been computed with the aid of Eq. (3) although the differences as compared with those obtained with Eq. (2) are not large.

## Results

Table I shows the results of 26 determinations of the dielectric constant of air corrected to normal temperature and pressure.

It shows the smallness of errors that are accidental in character. However, constant sources of error decrease the reliability of the result. The most important of these lies in the determination of the ratio of the stray to the effective capacity in the circuit. An error of 1 percent in this ratio will affect the final results for $(\epsilon-1)$ by 0.08 percent and this error is probably the limiting factor in the reliability of our data.

Table I. *Experimental values for dielectric constant of air at N.T.P.*

| | | |
|---|---|---|
| 1.0005899 | 1.0005899 | 1.0005895 |
| 1.0005899 | 1.0005895 | 1.0005892 |
| 1.0005897 | 1.0005902 | 1.0005900 |
| 1.0005893 | 1.0005910 | 1.0005895 |
| 1.0005902 | 1.0005888 | 1.0005897 |
| 1.0005903 | 1.0005903 | 1.0005902 |
| 1.0005903 | 1.0005888 | 1.0005901 |
| 1.0005907 | 1.0005889 | 1.0005900 |
| | 1.0005906 | 1.0005899 |
| | Average | 1.00058986 |
| | Average deviation from the mean | 0.00000043 |
| | Number of observations | 26 |

It may be noted that neither Eq. (2) or (3) contains a term for either the stray or the effective capacity of the circuit independently but only one for the ratio of these two quantities. Precision of measurement of these quantities is then only important insofar as it effects this ratio and an error in this ratio affects the value of $(\epsilon-1)$ only to the extent indicated above.

On this basis the value for the dielectric constant of air at normal pressure and 0°C (when the air is dried with $CaCl_2$ only and has had $CO_2$ removed by a solution of KOH) at 900,000 cycles per second is

$$1.00058986 \pm 0.00000050.$$

This value compares with 1.000591 found by Jordan, Broxon, and Walz[5] at 70,000 cycles per second in 1934 and with 1.000590 found by Boltzmann by static methods as long ago as 1874.

[5] A. R. Jordan, James W. Broxon and Frank C. Walz, Phys. Rev. **46**, 66 (1934).

# Low Pressure Arc Characteristics*

J. D. COBINE, *Graduate School of Engineering, Harvard University*

(Received July 29, 1935)

The results of an oscillographic study of the short a.c. arc between pure graphite electrodes show that the reignition potential is a linear function of the gas pressure for air, $N_2$, $O_2$, and $CO_2$. The relation is expressed by $V_r = A + Bp$. The constant $A$ follows the minimum sparking potential of the gas, increases with the rate of rise of the impressed voltage, and is independent of the arc current, except for $CO_2$. The constant $B$ depends on the gas, arc current, and rate of rise of impressed voltage. The reignition potential is closely related to the glow discharge and is shown to approach the sparking potential of the gas as the r.m.s. value of the current approaches zero. The factors that affect reignition are indicated, and a reasonable mechanism of the phenomenon is stated.

## INTRODUCTION

IN an experimental study of the factors affecting the reignition of a short a.c. arc between refractory electrodes, it has been shown[1] that for air the reignition potential increases with pressure, decreases with increasing current, and is markedly reduced by the addition of small amounts of low work-function impurities. The experimental results presented in this report indicate that for several gases there is a linear relationship between the reignition potential and the gas pressure in the case of a short arc between refractory electrodes.

## EXPERIMENTAL RESULTS

The apparatus and circuit are the same as previously described.[1] The power was supplied to the arc by a power transformer, and the r.m.s. arc current was limited to the desired value by a resistance in series with the arc and the transformer. Oscillograms were taken of the voltage across the gap by means of a magnetic oscillograph. The discharge investigated took place between the plane end surfaces of cylindrical electrodes of 6.4 mm diameter separated by a gap of 1 mm. The electrodes were placed under a large bell jar which was connected to suitable pumps and gauges. Data were taken for pressures above 10 cm of Hg. The results for pressures below this were not believed to be representative of the electrodes being studied, for the discharge was not confined to the electrodes; but covered all exposed parts of the electrode holders.

The electrodes were spectrographic carbons supplied by the National Carbon Company. According to the manufacturer these carbons were made of specially purified graphite having considerably less than 0.001 percent ash; examination had show the presence of silicon, calcium, magnesium, silver, and copper. It was stated that no carbon examined had shown the presence of more than two or three of these elements and some had shown the lines of carbon alone. As was pointed out previously,[1] electrodes of high purity exhibited very uniform reignition potentials from cycle to cycle.

The gases used were nitrogen, oxygen, and carbon dioxide of commercial purity, and air. The electrodes were cleaned and the gap adjusted at frequent intervals so that the degree of uniformity of surface and gap would be as high as possible. Each value of reignition potential is the average of about the first forty individual values taken from oscillograms. Deviations from the average were small. In the case of nitrogen it was found that no error was introduced when the deflections were taken directly from the oscillograph screen, that is, the time studied could cover several seconds.

There was practically no destruction of the electrodes when nitrogen was used. With oxygen there was considerable consumption of the graphite; an after-glow of short duration was noted between the electrodes when a current greater than 4 amperes was interrupted, indicating, possibly, sustained combustion. This effect was not noted in air, neither was the electrode consumption very large. The use of carbon dioxide resulted in considerable burning of the electrodes, probably due to the fact that $CO_2$ is

* Part of the experimental results of a thesis in partial fulfillment of the requirements for the degree of Doctor of Philosophy, California Institute of Technology, June 1934.

[1] S. S. Mackeown, J. D. Cobine and F. W. Bowden, Elect. Eng. **53**, 1081 (1934).

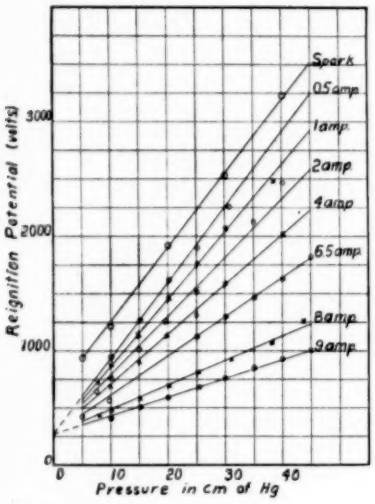


FIG. 1. Arc reignition potential for graphite in nitrogen. Transformer voltage = 2200. Parameter = r.m.s. arc current.

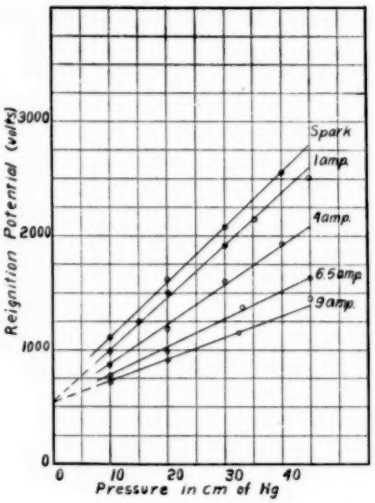


FIG. 2. Arc reignition potential for graphite in oxygen. Transformer voltage = 2200. Parameter = r.m.s. arc current.

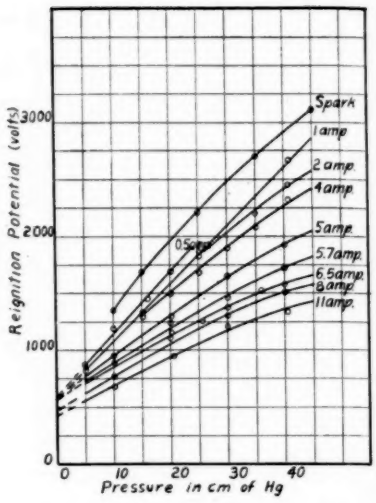


FIG. 3. Arc reignition potential for graphite in carbon dioxide. Transformer voltage = 2200. Parameter = r.m.s. arc current.

dissociated at high temperatures into O and CO.

The arc current is given in r.m.s. amperes, as indicated by an iron vane a.c. meter. Needless to say, the current wave is not sinusoidal; but has a very low or zero value during the period from arc extinction to reignition.

The experimental results showing the reignition potential as a function of gas pressure for different currents and gases are presented by the curves of Figs. 1 to 3. It is seen that in each case the reignition potential of the arc is a linear function of the pressure over a wide range of pressure; following a curve such as $V_r = A + Bp$. The dotted extension of the curves indicate the value of $A$. It should be noted that, except for carbon dioxide, each curve of the family has the same intercept at $p = 0$ for a given gas. The linear relation has been found to extend to atmospheric pressures for $O_2$, $N_2$, and air. In this case a higher transformer voltage was required so that the results may not be compared directly. Table I summarizes the constants for the two transformer voltages indicated.

TABLE I. *Reignition constants for graphite.*

| GAS | TRANSFORMER VOLTAGE | $A$ (volts) | MINIMUM SPARKING POTENTIAL[3] | $B$[1] (volts/cm Hg) 1 amp. | 6.5 amp. |
|---|---|---|---|---|---|
| Nitrogen | 2200 | 275 | 251–298 | 58.7 | 35.0 |
| | 3300 | 325 | | 43.0 | 27.0 |
| Air | 2200 | 325 | 327 | 75.0 | 40.0 |
| | 3300 | 375 | | 45.0 | 25.0 |
| Carbon Dioxide | 2200 | 425 to 600[2] | 420 | 51.3 | 25.0 |
| | 3300 | 400 to 600[2] | | 47.5 | 25.0 |
| Oxygen | 2200 | 550 | 439–455 | 46.0 | 25.0 |

[1] Determined for the straight line portion of the curves.
[2] Range for currents 11 to 1 amperes.
[3] J. J. Thomson, *Conduction of Electricity in Gases*, Vol. 2, third edition, p. 487.

The curves of Figs. 4 to 6 show the dependence of $V_r$, at constant gas pressure, upon the arc current. With nitrogen and with carbon dioxide there are two nonlinear regions in the range of current covered, just as was shown[1] to be true for air. One region for low currents, up to a little more than 6 amperes for $N_2$ and up to about 4 amperes for $CO_2$; and the other region involving higher currents. However in the case of oxygen the reignition potential decreases almost linearly with increasing current. The curves of Fig. 4 indicate that nitrogen as well as air[1] has, under the conditions of these experiments, a current at which the curves of constant pressure cross. When the current is increased above this value the highest pressures have the lowest $V_r$. This critical value of current marks the limit of the linear relationship between $V_r$ and pressure for air and nitrogen. This point was not reached for $O_2$ and $CO_2$. In all these curves the sparking potential is indicated for each pressure at zero current. The $V_r - I$ curves have the same shape

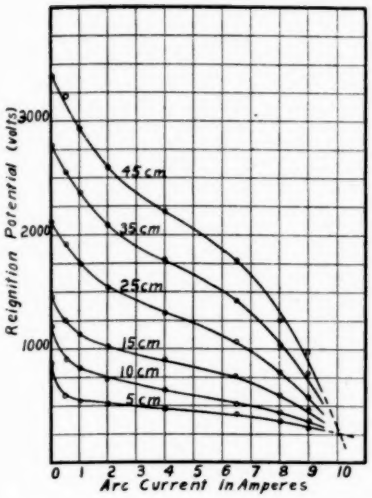


FIG. 4. Arc reignition potential for graphite in nitrogen as a function of the arc current. Parameter is gas pressure in centimeters of mercury.

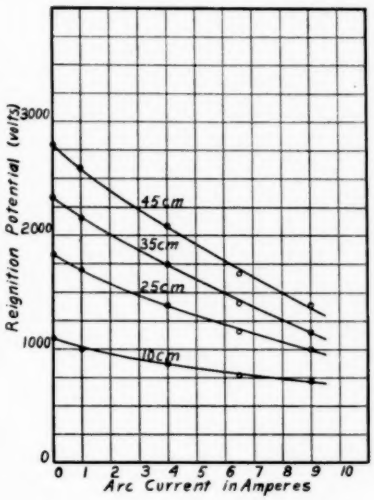


FIG. 5. Arc reignition potential for graphite in oxygen as a function of the arc current. The parameter is gas pressure in centimeters of mercury.

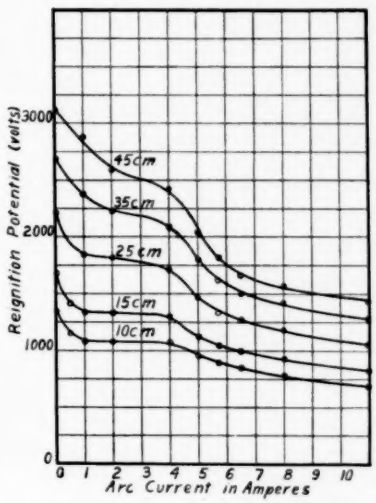


FIG. 6. Arc reignition potential for graphite in carbon dioxide as a function of arc current. The parameter is gas pressure in centimeters of mercury.

at the higher transformer voltage as for those shown in the figures.

Fig. 7 shows for oxygen and for nitrogen the dependence of the constant $B$ upon the arc current. For oxygen this constant is nearly a linear function of the arc current. For nitrogen the constant $B$ was found to depend upon $I$, $I^2$, and $I^3$ as shown by the equation: $B=67.7-9.9I+1.35I^2-0.1I^3$. For nitrogen the critical value of current at which the linear relationship between $V_r$ and pressure breaks down is indicated by the point at which $B$ becomes zero.

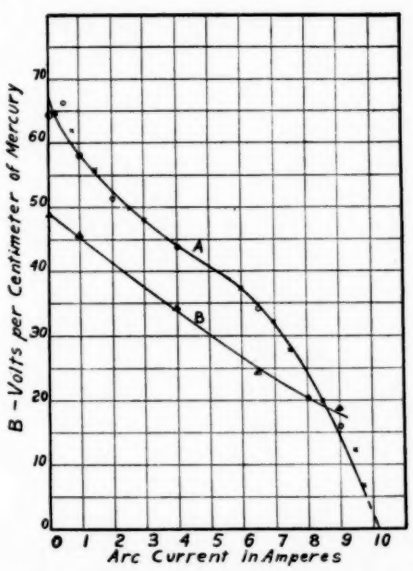


FIG. 7. The constant $B$ as a function of the arc current. $A$—circles, nitrogen, from Fig. 1; ×, nitrogen, from Fig. 4. $B$—oxygen, from Fig. 2.

The curves of Fig. 8 show the effect of the initial rate of rise of impressed, or transformer voltage, upon the reignition potential. It is evident that the effect of increasing $dV/dt$ in lowering $V_r$ is much less at low pressures than at the higher ones. The dotted portion of the curves is to indicate the sparking potential for the different pressures, i.e., the voltage approached as the rate of rise of impressed voltage following arc extinction is very small. The exact nature of the curve for the lower $dV/dt$ could not be determined. Although only a few points were possible, it is believed that the trend over this range is indicated.

## DISCUSSION

The experimental results establish the fact that for the material used, and a circuit of small reactance, the arc reignition potential is given, for a wide range of pressures, by $V_r=A+Bp$. Since the field is not uniform, the curves of $V_r$ *vs.* pressure would be expected to depart somewhat from linearity at the higher pressures; the marked departure of carbon dioxide may be due to chemical effects. The results summarized in Table I apply, of course, only to the material and

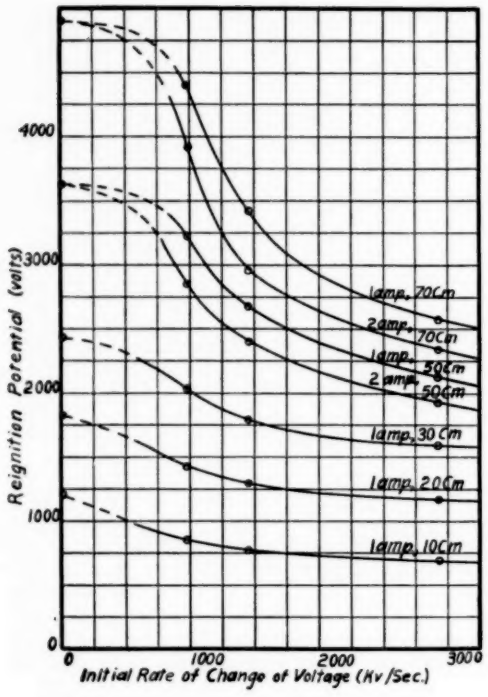


FIG. 8. Arc reignition potential of graphite in nitrogen as a function of the initial rate of rise of impressed voltage. The parameters are arc current and gas pressure.

shape of electrodes used in this work; however they are of interest in that they indicate the variables present. The constant $A$ follows the minimum sparking potential and quite probably the normal cathode drop, although this is not known for the materials and gas condition of the surfaces of the electrodes used. The fact that $A$ is larger than the minimum sparking potential for oxygen and carbon dioxide may be due to gas impurities which are known to greatly affect the minimum sparking potential. $A$ depends upon the rate of rise of impressed voltage, seeming to increase with this variable. This constant is independent of arc current except for carbon dioxide for which it increases as the current is decreased. The constant $B$ depends on the nature of the gas, the current during the previous half-cycle of arcing, the frequency or $dV/dt$, and quite probably on the gap and electrode material. It will be noted that the curves of $B$ *vs.* $I$ have the same form as the curves of $V_r$ *vs.* $I$ for the same gas. It is obvious that with the same gas and the same electrode materials this should be the case. Since the experimental results were quite uniform, a knowledge of the value of the constant $A$ and of the equation for $B$ makes it possible to predict the value of the arc reignition potential over a considerable range of current and pressure. However, since the constants of the $B$ *vs.* $I$ curve are purely empirical they give no clue to the particular factors involved.

The linear relation between $V_r$ and pressure, for the refractory electrodes used in nitrogen and in air, breaks down at a critical current for which the reignition potential is independent of pressure and has a value about equal to or somewhat greater than the normal cathode drop of a glow discharge. For currents larger than this value the reignition potential for graphite in air[1] decreases with increase in pressure; and the curves of Fig. 4 indicate that this is probably true for nitrogen. The critical current indicates, probably, the approximate point at which the arc is reestablished without a preliminary glow; the residual thermionic emission due to heating by the previous half-cycle of arcing being sufficient to restrike the arc with a voltage only a little above the arc burning voltage.

For arc currents less than the critical value, a different situation exists. In this case thermionic emission, although of considerable magnitude for currents near the critical, operates in conjunction with a number of other factors in effecting reignition; all of which vary in relative importance with the state of the electrode surface, such as irregularities, impurities, adsorbed gases, etc. and also depend upon the nature of the gas, the pressure, etc. For the lower currents the reignition potential for carbon, if pure, is approximately that of the relatively low melting point metals. Small amounts of low work function impurities will cause the reignition potential to vary from cycle to cycle, and larger amounts will lower it markedly.[1] The coefficient of heat conductivity, which probably varies with carbons of different manufacture, will also affect reignition by changing the rate of cooling of the surfaces after current zero.

The effect of increasing the rate of rise of impressed voltage (increasing the transformer voltage at constant frequency—assuming a sine wave) in lowering the reignition potential is probably due to the fact that less time is available for the cooling of the electrode surface. Also, less time is available for the deionizing of the gap.

Since the reactance of the transformer was fixed and the current adjusted by varying a series resistance, the power factor of the circuit external to the arc itself, decreased as the r.m.s. value of arc current was increased. For the case where the effect of this constant circuit reactance would be greatest, i.e., for a large arc current (variable resistance small) and low pressure (where the required voltage for reignition is small), it was found that the angle of lag was less than 6° and in this case reignition occurred 3.5° after zero current. Under these conditions approximately 60 percent of the reignition potential was reached with a very high rate of rise of voltage upon zero current and the balance at the relatively low rate of rise of the impressed voltage.[2] For lower currents this effect was negligible because of the relatively high value of series resistance required.

The results would seem to indicate that Paschen's law can be extended to the reignition potential of the low pressure, short arc between refractory electrodes. It is necessary in this case, however, to specify the arc current since the constant $B$ is different for each value of current. Recent glow discharge studies[3, 4] have indicated that Paschen's law holds for the reignition and burning potentials of the glow, and that the glow reignition potential is reduced when the rate of rise of applied voltage is increased.[5] Evidently the processes involved in the arc and glow reignitions are similar.

It is probable that the early stages of the arc reignition are similar to the glow discharge reignition. The difference between the two at the instant of zero current is due to the difference in the amount of ionization in the two cases. In the case of the arc a cathode spot must be formed. Hence the process of reignition must include a gaseous state in which conditions are such that this cathode spot may be formed. In previous papers[1, 6] it has been suggested that the reignition of an arc consists in the establishing of a "normal" glow discharge which quickly becomes "abnormal," the gap voltage and electrode current density increasing as the applied voltage increases. Then at some critical voltage the transition to the arc is effected. This transition is known to take place very quickly if the conditions are favorable. The theory[1, 6] proposes that the arc cathode spot forms as the result of the cumulative effects of positive ion bombardment, high electric field and high temperature on localized emission at the new cathode. Essentially the same suggestion has been made by Newman.[7]

The assumption of a preliminary abnormal glow provides a gas condition such that the transition voltage would be expected to vary just as it has been found to do in these experiments. First, there will be a considerable amount of positive ion bombardment and second, the positive ion space charge of the glow will produce a high electric field at the cathode. The electron emission due to positive ion bombardment will depend upon the gas condition of the surface and the ion energy. The electrode surface being far from a true plane because of submicroscopic points, projecting crystal corners, etc., will be subjected to very high local electric fields. In these experiments the applied voltage will reach a value equal to the normal cathode drop in a time of one to three hundred microseconds (50 cycles, 6600 and 2200 volts r.m.s.). When this point is reached the glow will start if the factors affecting the deionization of the gap have not been too active. These factors are recombination, diffusion, and reduction of thermionic emission of the electrodes due to cooling.

It has been shown[8] that in a few microseconds a short, turbulent, heavy current arc will recover a dielectric strength of several hundred volts per centimeter. The remainder of its dielectric strength is recovered at a very much lower rate. The results of Browne's work are not directly applicable to the present case because they are for a longer arc, and an arc for which the turbulence and current were very much greater. However, the effects of turbulence and current upon the dielectric strength are opposite to each other. Reducing the turbulence will reduce the dielectric strength of the gas at a given instant after zero current and reducing the current will increase the dielectric strength. Thus a rather

[2] J. Slepian, *Conduction of Electricity in Gases*, Westinghouse Course Number 38.
[3] H. Fricke, Physik. Zeits. **34**, 168 (1933).
[4] J. Thomson, Phil. Mag. **18**, 696 (1934).
[5] W. Spielhagen, Physik. Zeits. **34**, 164 (1933).
[6] S. S. Mackeown, Elect. Eng. **51**, 386 (1932).
[7] F. H. Newman, Phil. Mag. **18**, 365 (1934).
[8] T. E. Browne, Jr., Physics **5**, 103 (1934).

rough comparison of results may be made. It is evident that the dielectric strength indicated by the Browne results is exceeded early in the reignition period in the present experiments, so that a glow discharge will result from the electric failure of the ionized gas. Further, it is interesting to note that the order of decreasing dielectric strength after one hundred microseconds is carbon dioxide, oxygen, and air, and that the order of decreasing reignition potential is the same. This is as it should be, since the gases having the highest recovered dielectric strength will require a higher voltage to initiate the preliminary discharge and will, no doubt, require a higher transition voltage as has been observed. The order indicated does not hold after thermionic emission has become a considerable factor since the experimental conditions have then been entirely altered. When the arc current is very small, it is necessary for the voltage to rise to the sparking value before the arc is struck. When the current is large, the thermionic emission for the graphite electrodes is high and the arc forms at a relatively low voltage without the preliminary glow.

Recently published results would seem to justify the belief that a high electric field is an important factor in the glow-arc transition. The emission from a heated cathode may be increased by a factor of 10 by the action of a positive ion space-charge field.[9] It has been shown by Beams[10] that cold emission will start from impure mercury surfaces with *average* fields of $3.5\times10^5$ volts per centimeter and for pure mercury at $1.8\times10^6$ volts per centimeter; the latter value being considerably less than that required by the Fowler-Nordheim wave mechanics formula. It has been recently pointed out that emission at relatively low calculated fields could exist because of points on the surface of the electrodes.[11, 12, 13] In addition to the natural irregularities of the electrode surfaces, there is the effect produced by particles deposited upon the surface from the other electrode, or from other surfaces.[14] This effect will introduce an important change in the state of purity of the electrode surface as well as affecting the field, if the arc is studied at pressures below about 7 cm of mercury. Under this condition the arc cannot well be confined to the electrodes under study; but spreads to all exposed surfaces at different potentials and will therefore cause relatively large amounts of foreign material to be deposited upon the electrodes. Spectrograms recently taken at Harvard of the pure carbon before and after being arced at low pressures (5 cm of mercury) indicated that a considerable amount of copper and other impurities had been deposited on the electrode surfaces, the nearest source being about twenty times the gap length from the gap.

Another effect which cannot be overlooked is that which may result due to chemical reactions which may take place between the hot electrode vapor at and near the surface and the gas during the period of reignition. There is also the effect of dissociation of the gas and the effect of the ionization potential of any electrode vapor which may remain in the arc space after zero current. Just as all of these factors acting together affect the ionization or deionization of the gap, so will they lower or raise the voltage required to reignite the arc. A satisfactory theory of arc reignition must take into account all the factors outlined since they contribute in varying degrees to the reignition.

The experimental results show the effect of increased current which increases the residual thermionic emission and the volume of ionized gas, and therefore lowers the reignition potential as required by the theory. A similar effect is due to increasing the rate of rise of impressed voltage. The close relation between the reignition potential and the striking and burning voltages of the glow discharge also points to the correctness of the assumption of a preliminary abnormal glow.

## Acknowledgment

The writer wishes to thank the electrical engineering staff of the California Institute of Technology for the facilities provided and assistance rendered during this research. Many helpful suggestions were made by Dean H. E. Clifford of the Graduate School of Engineering of Harvard University during the preparation of the manuscript.

[9] C. G. Found, Phys. Rev. **45**, 526 (1934).
[10] J. W. Beams, Phys. Rev. **44**, 803 (1933).
[11] A. J. Dempster, Phys. Rev. **46**, 728 (1934).
[12] C. C. Chambers, J. Frank. Inst. **218**, 463 (1934).
[13] L. Tonks, Elect. Eng. **53**, 239 (1934).
[14] W. H. Bennett, Phys. Rev. **44**, 859 (1933).

# Quasi-Holonomic Dynamical Systems

GABRIEL KRON, *Engineering General Department, General Electric Company, Schenectady*

(Received September 18, 1935)

When a dynamical system contains both *mechanical* and *electromagnetic* energies, the dynamical equations assume special forms. It is shown that when the dynamical system is nonholonomic, or when it is holonomic but is referred to nonholonomic reference axes, the equations still have the same form as those of a holonomic dynamical system and can be solved as such without any auxiliary equations. Following a suggestion of Lorentz, such dynamical systems are called here "quasi-holonomic" dynamical systems. Because of the special form of the equations, their development in tensor notation allows the introduction of concepts from the absolute calculus that apparently have not been hitherto applied in classical dynamics. It is shown here that the equation of motion of quasi-holonomic systems in tensor symbolism may be interpreted as specifying the trajectory of a point moving in an ($n$-dimensional) *non-Riemannian* space, referred to holonomic or nonholonomic reference axes. When the field equations of Maxwell in tensor form are formulated for the electromagnetic part of the quasi-holonomic system, it is found that the electromagnetic field must be expressed in terms of its components along an orthogonal ennuple, the field lying in a local Euclidean space, tangent to the underlying non-Riemannian space. *By the use of such a nonholonomic reference frame the dynamical equations of Lagrange and the field equations of Maxwell are correlated.* The equations show *formal* analogies with those of the five-dimensional unified field theory of relativity dynamics, where *combined* electromagnetic and gravitational systems are analyzed. Practical examples of quasi-holonomic systems are the numerous types of rotating electrical machinery used in industry.

## INTRODUCTION

THE close analogy between the equations of performance of rotating electrical machinery and the equations of the differential geometry of abstract spaces (developed by the requirements of modern electrodynamics) has been pointed out elsewhere,[1] illustrated with numerous examples. The key to the analogy was the recognition that the various *interconnections* of electrical networks and machines represent a group of transformations that leave several quadratic differential forms (stored magnetic energy, etc.) invariant.

The performance of electromechanical systems may be represented by the motion of a point in an $n$-dimensional representative space. The following types of spaces may be distinguished:

(1) When the magnetic paths, the reference axes and the electric conductors are all stationary, the space is a Euclidean space, having neither curvature nor torsion.

(2) When the magnetic paths alone move, the space is a Riemannian space, having curvature, but no torsion.

(3) When the reference axes alone move, the space is a space of distant parallelism, having torsion, but no curvature.

(4) When the electric conductors alone move, the space is a group space, having both curvature and torsion, but the components of the curvature and torsion tensors are all constants, satisfying the fundamental structural relations of the theory of finite, continuous groups.

(5) When the magnetic paths, the reference axes and the conductors all move with different velocities, the space is a general non-Riemannian space (metric space), with its curvature and torsion tensors satisfying the generalized identities of Bianchi.

When a cross section is made of any rotating machine during acceleration, at any one instant the current-density and flux-density waves are sinusoidally distributed in space. These space vectors may be considered as projections of vectors upon various two-dimensional planes, the vectors lying in a *tangent* $n$-dimensional Euclidean space and originating at the center of the rotor. That is, an electrical machine offers a visualization of vectors of the electromagnetic field lying in a *centered* local space, tangent to the underlying non-Riemannian (metric) space and moving along a trajectory in the underlying space. The translatory motion of the tangent space is accompanied by a rotation and slippage.

In this paper the recently developed theory of nonholonomic reference frames[2] is used to put

[1] Kron, J. Math. and Phys., 103–194 (1934); Kron, Gen. Elec. Rev. A serial beginning in the April, 1935 issue.

[2] Schouten and Struik, *Einführung in die neueren Methoden der Differentialgeometrie* (1935). The notation and ter-

the previously given dynamical equations in a more general form, so as to correlate them with the field equations of Maxwell, when they are applied to the electromagnetic part of a *combined* mechanical and electrical system. Just as in the unified field theory, *no room can be found within the framework of a Riemannian space to represent the electromagnetic field in the presence of a mechanical system*. Hence even for *classical* dynamical systems the representative space *must* be non-Riemannian, to take care of the interactions between mechanical and electrical systems.

## I. Nonholonomic Reference Systems

(a) Let the geometric form of the equation of motion of a nonconservative *holonomic* dynamical system, expressed along holonomic reference axes, be given as

$$a_{km}\frac{d^2x^m}{dt^2}+[mn,\,k]\frac{dx^m}{dt}\frac{dx^n}{dt}=f_k, \tag{1}$$

where $a_{km}$ and $f_k$ are functions of $x^m$. The holonomic form of the Christoffel symbol is defined as

$$[mn,\,k]=\overset{h}{\frac{1}{2}}\left(\frac{\partial a_{nk}}{\partial x^m}+\frac{\partial a_{mk}}{\partial x^n}-\frac{\partial a_{mn}}{\partial x^k}\right). \tag{2}$$

Since frictional forces do not play any fundamental part in the analysis, they will be omitted for the time being to make the equations simpler.

(b) Let $dx^m$ be replaced by a new set of differentials $dx^\alpha$

$$dx^m=A_\alpha{}^m dx^\alpha \tag{3}$$

by a nonholonomic relation, where $A_\alpha{}^m$ is a function of the old variables $x^m$ and where

$$A_\alpha{}^m\neq\partial x^m/\partial x^\alpha, \tag{4}$$

That is the relation (3) is not integrable. As a consequence

$$\partial A_m{}^\alpha/\partial x^n\neq\partial A_n{}^\alpha/\partial x^m. \tag{5}$$

(c) By replacing also $a_{km}$ by $a_{\gamma\alpha}A_k{}^\gamma A_m{}^\alpha$ etc., the equation of motion expressed along the new *nonholonomic* reference system becomes

$$a_{\gamma\alpha}\frac{d^2x^\alpha}{dt^2}+[\alpha\beta,\,\gamma]\frac{dx^\alpha}{dt}\frac{dx^\beta}{dt}=f_\gamma, \tag{6}$$

where the nonholonomic form of the Christoffel symbol is defined as

$$[\alpha\beta,\,\gamma]=\tfrac{1}{2}(\partial a_{\gamma\beta}/\partial x^\alpha+\partial a_{\gamma\alpha}/\partial x^\beta-\partial a_{\alpha\beta}/\partial x^\gamma)+\Omega_{\gamma\beta,\,\alpha}+\Omega_{\gamma\alpha,\,\beta}-\Omega_{\alpha\beta,\,\gamma} \tag{7}$$

where $\Omega_{\alpha\beta}{}^\delta$ is the "*nonholonomic object*"

$$\Omega_{\alpha\beta}{}^\delta=\tfrac{1}{2}A_\alpha{}^m A_\beta{}^n(\partial A_m{}^\delta/\partial x^n-\partial A_n{}^\delta/\partial x^m), \tag{8}$$

skew-symmetric in $\alpha$ and $\beta$, with a transformation formula

$$\Omega_{\alpha\beta}{}^\delta=\Omega_{\sigma\epsilon}{}^\omega A_\alpha{}^\sigma A_\beta{}^\epsilon A_\omega{}^\delta+A_\alpha{}^\sigma A_\beta{}^\epsilon(\partial A_\sigma{}^\delta/\partial x^\epsilon-\partial A_\epsilon{}^\delta/\partial x^\sigma), \tag{9}$$

and where $\partial a_{\gamma\beta}/\partial x^\alpha$ is written for $(\partial a_{\gamma\beta}/\partial x^n)A_\alpha{}^n$. In analogy to the Christoffel symbol the upper index is lowered as follows,

$$\Omega_{\alpha\beta,\,\gamma}=\Omega_{\alpha\beta}{}^\delta a_{\delta\gamma} \tag{10}$$

to simplify the form of the equations.

*In the $n$ dynamical Eqs. (6) there are $2n$ unknowns;* the $n$ old variables $x^k$ and the $n$ new differentials $dx^\gamma$, since the nonholonomic transformation (3) replaces only the old differentials, but not the old variables. The other $n$ equations are the $n$ equations of transformations (3).

## II. Quasi-Holonomic Transformations

(a) Let the $n$ old variables $x^k$ be divided into two groups containing $p$ and $n-p$ variables and let the following assumption be made:

(1) The components of the metric tensor $a_{km}$ are functions only of the variables of the second group.

Also let a group of transformations be introduced with the following additional assumptions:

(2) The variables of the second group remain unchanged under the transformations to be considered, hence their differentials also remain unchanged.

(3) The differentials of the variables of the first group are replaced by new differentials, in such a way that the elements in their matrix of transformation are functions only of the variables of the second group.

(A more general case is to allow the variables of the second group to transform among themselves, instead of keeping them unchanged.)

(b) Such a group of transformations occur in electrical machinery in going from reference axes moving with the conductors to axes moving with a different speed than the conductors. The first

---

minology of this book is followed, except for the order of the indices of the various geometric objects which is kept the same as in Schouten's *Der Ricci-Kalkül*.

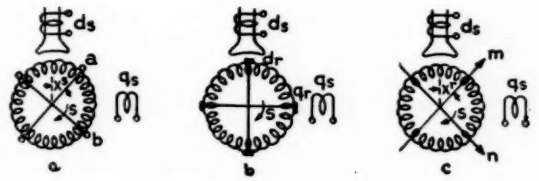


FIG. 1. Rotating machines and their reference axes. *a*, holonomic, Riemannian; *b*, holonomic, non-Riemannian; *c*, nonholonomic, non-Riemannian. Open circle = reference axis, connected to moving conductors; solid rectangle = reference axis, stationary in space; arrow = reference axis, moving at different speed than conductors.

group of variables are the *electric charges* passing through the various conductors, the second group of variables are the *displacements* of:

(1) The moving reference axes,
(2) the moving rotor conductors,
(3) the moving magnetic paths.

A representative example is the transformation (Fig. 1, Section V)

$$\begin{aligned} dx^{d_s} &= dx^{d_s}, \\ dx^a &= \cos x^s dx^{d_r} + \sin x^s dx^{q_r}, \\ dx^b &= -\sin x^s dx^{d_r} + \cos x^s dx^{q_r}, \\ dx^{q_s} &= dx^{q_s}, \\ dx^s &= dx^s, \end{aligned} \tag{11}$$

with a matrix of transformation

$A_\alpha^m = b$

| $m$ \ $\alpha$ | $d_s$ | $d_r$ | $q_r$ | $q_s$ | $s$ |
|---|---|---|---|---|---|
| $d_s$ | 1 | 0 | 0 | 0 | 0 |
| $a$ | 0 | $\cos x^s$ | $\sin x^s$ | 0 | 0 |
| $b$ | 0 | $-\sin x^s$ | $\cos x^s$ | 0 | 0 |
| $q_s$ | 0 | 0 | 0 | 1 | 0 |
| $s$ | 0 | 0 | 0 | 0 | 1 |

(12)

where $x^s$ is the geometrical variable representing the rotor displacement.

In general, quasi-holonomic transformations may occur in such dynamical systems in which both electrical and geometrical variables exist, since then $a_{mn}$ is a function only of the geometrical variables.

Quasi-holonomic transformations are also used in the five-dimensional projective relativity[3] where also:

(1) The metric tensor is independent of the electrical coordinates,
(2) in changing to a new reference system the space-time coordinates (geometrical variables) remain unchanged, or are transformed only among themselves,
(3) a nonholonomic relation exists between the electrical axes,
(4) the transformation matrix is independent of the electrical coordinates.

(In the original five-dimensional unified field theory of Kaluza-Klein, the above cylindrical properties of the electrical variables also are assumed. However, the relation between the old and the new electrical variables is holonomic. For that reason the trajectory of the point in their theory is Riemannian.)

(c) After the transformation is made, the $n$ new nonholonomic equations contain only $n$ unknowns, the $n-p$ variables of the second group and the $p$ new differentials of the first group. In the new set of equations the old variables appear as new variables and the new equations can be solved as if they were holonomic. Hence, the new equations may be considered to describe the motion of a holonomic system.

## III. The Equation of Motion

(a) Since the nonholonomic object $\Omega_{\alpha\beta,\ \gamma}$ (and every other geometric object) is a function now only of the variables of the *new* reference system, let its components be considered to form the components of another geometric object $S_{\alpha\beta\gamma}$, a *tensor* of rank three, skew-symmetric in $\alpha$ and $\beta$

$$\Omega_{\alpha\beta,\ \gamma} \overset{*}{=} -S_{\alpha\beta\gamma}. \tag{13}$$

In other words, let a new tensor $S_{\alpha\beta\gamma}$ be introduced whose components are identical (but negative) with the components of $\Omega_{\alpha\beta,\ \gamma}$ for this one reference system only, but different for other reference systems, since the transformation formula of $S_{\alpha\beta\gamma}$ is

$$S_{\alpha\beta\gamma} = S_{\iota\sigma\omega} A_\alpha{}^\iota A_\beta{}^\sigma A_\gamma{}^\omega, \tag{14}$$

while that of $\Omega_{\alpha\beta,\ \gamma}$ is the one given in Eq. (9).

(b) Similarly, in this new reference system let $[\alpha\beta, \gamma]$ of Eq. (7) be replaced by another geometric object $\Gamma_{\alpha\beta,\ \gamma}$ containing $S_{\alpha\beta\gamma}$ instead of $\Omega_{\alpha\beta,\ \gamma}$

$$\Gamma_{\alpha\beta,\ \gamma} \overset{h}{=} [\alpha\beta, \gamma] + S_{\alpha\beta\gamma} - S_{\gamma\beta\alpha} - S_{\gamma\alpha\beta}, \tag{15}$$

[3] Banesh Hoffman, Phys. Rev. **43**, 615 (1933).

where now $[\alpha\beta, \gamma]$ is the holonomic Christoffel symbol defined in Eq. (2), since $(\partial a_{\gamma\beta}/\partial x^n)A_\alpha{}^n$ can be written in quasi-holonomic systems as $\partial a_{\gamma\beta}/\partial x^\alpha$.

(c) If a new tensor is introduced to replace the three tensors

$$T_{\alpha\beta\gamma} = S_{\alpha\beta\gamma} - S_{\gamma\beta\alpha} - S_{\gamma\alpha\beta}, \tag{16}$$

then the new affine connection is defined as

$$\Gamma_{\alpha\beta,\ \gamma} = [\alpha\beta, \gamma] + T_{\alpha\beta\gamma} \tag{17}$$

and the equation of motion (6) becomes for the new quasi-holonomic system:[4]

$$a_{\gamma\alpha}\frac{d^2x^\alpha}{dt^2} + \Gamma_{\alpha\beta,\ \gamma}\frac{dx^\alpha}{dt}\frac{dx^\beta}{dt} = f_\gamma, \tag{18}$$

$$a_{\gamma\alpha}\ddot{x}^\alpha + [\alpha\beta, \gamma]\dot{x}^\alpha\dot{x}^\beta + T_{\alpha\beta\gamma}\dot{x}^\alpha\dot{x}^\beta = f_\gamma. \tag{19}$$

(d) If the asymmetrical part of $\Gamma_{\alpha\beta,\ \gamma}$ is taken in Eq. (15) then

$$\Gamma_{[\alpha\beta],\ \gamma} \overset{h}{=} S_{\alpha\beta\gamma}, \tag{20}$$

since $S_{\alpha\beta\gamma}$ is skew-symmetric in its first two indices. Hence, the tensor $S_{\alpha\beta\gamma}$ is identified as the torsion tensor used by Schouten and Struik. It also follows that the restriction

$$\nabla_\alpha a_{\beta\gamma} = 0 \tag{21}$$

exists, since the affine connection in Eq. (15) does not contain the additional tensor $Q_\alpha{}^{\beta\gamma} = \nabla_\alpha a^{\beta\gamma}$. Hence the non-Riemannian space of quasi-holonomic systems is a *metric* space.

(e) The effect of the frictional forces is to add an additional term $R_{\gamma\alpha}\dot{x}^\alpha$, so that the complete equation of motion of the quasi-holonomic system is

$$a_{\gamma\alpha}\ddot{x}^\alpha + \Gamma_{\alpha\beta,\ \gamma}\dot{x}^\alpha\dot{x}^\beta + R_{\gamma\alpha}\dot{x}^\alpha = f_\gamma. \tag{22}$$

Such generalized trajectories have already been used in differential geometry.[5]

(f) In terms of *tensors* only, the equation of motion is

$$\delta\dot{x}^\gamma/\delta t + R_\alpha{}^\gamma\dot{x}^\alpha = f^\gamma. \tag{23}$$

(g) In going over from this reference system to a third system by a quasi-holonomic transformation, the equation of the new system is the same as that of the old, Eqs. (18) or (19), except now $[\alpha\beta, \gamma]$ will be the nonholonomic Christoffel symbol defined in Eq. (7). Hence the affine connection $\Gamma_{\alpha\beta,\ \gamma}$ of this new system may be considered as being *nonholonomic and non-Riemannian.*

## IV. The Equations of Motion of Lagrange

(a) If, in the holonomic form of the Lagrangian equation of motion

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{x}^k} - \frac{dT}{\partial x^k} = f_k$$

$T$ is replaced by $(1/2)a_{mn}\dot{x}^m\dot{x}^n$, the resulting equation is

$$a_{km}\frac{d^2x^m}{dt^2} + \left(\frac{\partial a_{mk}}{\partial x^n} - \frac{1}{2}\frac{\partial a_{mn}}{\partial x^k}\right)\dot{x}^m\dot{x}^n = f_k. \tag{24}$$

It should be noted that the term in the parenthesis is not the Christoffel symbol. However, it can be replaced by it since

$$\frac{\partial a_{mk}}{\partial x^n}\dot{x}^m\dot{x}^n = \frac{1}{2}\left(\frac{\partial a_{mk}}{\partial x^n} + \frac{\partial a_{nk}}{\partial x^m}\right)\dot{x}^m\dot{x}^n.$$

(b) If a nonholonomic transformation is introduced in Eq. (24) instead of in Eq. (1), just as it was done in Section I, and $a_{km}$ is replaced by $a_{\gamma\alpha}A_k{}^\gamma A_m{}^\alpha$ etc., the resulting equation of motion has a different form from Eq. (6)

$$a_{\gamma\alpha}\ddot{x}^\alpha + \left[\left(\frac{\partial a_{\alpha\gamma}}{\partial x^\beta} - \frac{1}{2}\frac{\partial a_{\alpha\beta}}{\partial x^\gamma}\right) + A_\gamma{}^k A_\alpha{}^n\left(\frac{\partial A_k{}^\delta}{\partial x^n} - \frac{\partial A_n{}^\delta}{\partial x^k}\right)a_{\delta\beta}\right]\dot{x}^\alpha\dot{x}^\beta = f_\gamma. \tag{25}$$

If $a_{\gamma\alpha}\dot{x}^\alpha$ is replaced by $\partial T/\partial\dot{x}^\gamma$, the resulting equation of motion is

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}^\gamma}\right) - \frac{\partial T}{\partial x^\gamma} + \frac{\partial T}{\partial \dot{x}^\delta}A_\gamma{}^k A_\alpha{}^n\left(\frac{\partial A_k{}^\delta}{\partial x^n} - \frac{\partial A_n{}^\delta}{\partial x^k}\right)\dot{x}^\alpha = f_\gamma. \tag{26}$$

This is the Boltzmann-Hamel form of the equation of motion valid for nonholonomic

[4] This equation with identical connection has been used for instance by Infeld, Zeits. f. Physik **50**, 137 (1928), or by Banesh Hoffmann, reference 3, Eq. (13), etc.

[5] Veblen and Whitehead, "The Foundations of Differential Geometry," Cambridge Tracts in Math. and Math. Phys., No. 29, p. 72 (1932), Eq. 14.1.

reference axes as given by Whittaker, Schaeffer, etc.[6]

(c) Eq. (25) can also be written for quasi-holonomic systems as

$$a_{\gamma\alpha}\ddot{x}^{\alpha}+[\alpha\beta,\ \gamma]\dot{x}^{\alpha}\dot{x}^{\beta}-2S_{\gamma\alpha\beta}\dot{x}^{\alpha}\dot{x}^{\beta}=f_{\gamma}. \quad (27)$$

If a new tensor is introduced

$$P_{\alpha\beta\gamma}=-2S_{\gamma\alpha\beta} \quad (28)$$

skew-symmetric in its first and third indices, the *equation of motion* becomes

$$a_{\gamma\alpha}\ddot{x}^{\alpha}+[\alpha\beta,\ \gamma]\dot{x}^{\alpha}\dot{x}^{\beta}+P_{\alpha\beta\gamma}\dot{x}^{\alpha}\dot{x}^{\beta}=f_{\gamma}. \quad (29)$$

If a new connection is defined

$$\Lambda_{\alpha\beta,\ \gamma}\overset{h}{=}[\alpha\beta,\ \gamma]+P_{\alpha\beta\gamma}\overset{h}{=}[\alpha\beta,\ \gamma]-2S_{\gamma\alpha\beta} \quad (30)$$

the dynamical form of the *equaiion of motion* is

$$a_{\gamma\alpha}\ddot{x}^{\alpha}+\Lambda_{\alpha\beta,\ \gamma}\dot{x}^{\alpha}\dot{x}^{\beta}=f_{\gamma}. \quad (31)$$

Both forms (31) and (18) of the equation of motion are given (without $f_{\gamma}$) in the unified field theory by the various writers as representing the trajectory of a charged particle. However they are not derived there from the equation of motion of Lagrange, but are set up from various geometrical, dynamical and philosophical considerations by a cut and try process.

(d) Although the two forms of the equation of motion (18) and (31) have different affine connections (the geometric connection (15) and the dynamical connection (30)) *both equations give identical answers*, since from Eqs. (16) and (19)

1. $S_{\alpha\beta\gamma}\dot{x}^{\alpha}\dot{x}^{\beta}=0$, $S_{\alpha\beta\gamma}$ being skew-symmetric in its first two indices.[7]

2. $(-S_{\gamma\alpha\beta}-S_{\gamma\beta\alpha})\dot{x}^{\alpha}\dot{x}^{\beta}=-2S_{\gamma\alpha\beta}\dot{x}^{\alpha}\dot{x}^{\beta}$ which is the same as Eq. (27).

However, in equations where the connections are not multiplied by $\dot{x}^{\alpha}\dot{x}^{\beta}$, each of them appears to give a different answer. Such a case occurs for instance in trying to set up the field equations of Maxwell.

(e) The tensor $P_{\gamma\alpha\beta}$ is defined in Eq. (25) as

$$P_{\alpha\beta\gamma}=A_{\gamma}{}^{k}A_{\alpha}{}^{n}(\partial A_{k}{}^{\delta}/\partial x^{n}-\partial A_{n}{}^{\delta}/\partial x^{k})a_{\delta\beta}. \quad (32)$$

Let $A_{k}{}^{\delta}$ be considered as a set of vectors $A_{k}{}^{(\delta)}$. Since the rotation-coefficients of Ricci are defined as

$$\gamma_{\gamma\alpha\beta}\overset{*}{=}A_{(\gamma)}{}^{k}A_{(\alpha)}{}^{n}\overset{R}{\nabla}_{n}A_{k}{}^{(\delta)}a_{\delta\beta} \quad (33)$$

(where the covariant derivative is taken with respect to the Christoffel symbol) the tensor $P_{\gamma\alpha\beta}$ can be expressed in terms of them[8] as

$$P_{\gamma\alpha\beta}\overset{*}{=}\gamma_{\gamma\alpha\beta}-\gamma_{\alpha\gamma\beta}. \quad (34)$$

Also if the Christoffel symbol is zero

$$\Gamma_{\alpha\beta,\ \gamma}\overset{*}{=}\gamma_{\alpha\beta\gamma}. \quad (35)$$

## V. Rotating Electrical Machinery

(a) Among the numerous types of electrical machines the performances of those machines only are given by Eq. (1) (holonomic, Riemannian) where the reference axes are rigidly connected to the stationary or moving conductors. (Fig. 1a.)

Its metric tensor $a_{mn}$ represents all the self and mutual inductances of the various windings as a function of the rotor displacement angle $x^{s}$ and the moment of inertia of the rotor $L$. The resistance tensor $R_{mn}$ (Eq. (22)) represents the resistances of the various windings. The metric tensor is

| $n$ / $m$ | $d_s$ | $a$ | $b$ | $q_s$ | $s$ |
|---|---|---|---|---|---|
| $d_s$ | $L_{d_s}$ | $M_d\cos x^s$ | $-M_d\sin x^s$ | 0 | 0 |
| $a$ | $M_d\cos x^s$ | $L_{d_r}\cos^2 x^s+L_{q_r}\sin^2 x^s$ | $(L_{q_r}-L_{d_r})\sin x^s\cos x^s$ | $M_q\sin x^s$ | 0 |
| $a_{mn}=b$ | $-M_d\sin x^s$ | $(L_{q_r}-L_{d_r})\sin x^s\cos x^s$ | $L_{d_r}\sin^2 x^s+L_{q_r}\cos^2 x^s$ | $M_q\cos x^s$ | 0 |
| $q_s$ | 0 | $M_q\sin x^s$ | $M_q\cos x^s$ | $L_{q_s}$ | 0 |
| $s$ | 0 | 0 | 0 | 0 | $L$ |

(36)

[6] Whittaker, *Analytical Dynamics*, p. 43.

[7] The fact that this term drops out of the equation of motion is also mentioned by Mandel, Zeits. f. Physik **56**, 838 (1929).

[8] McVittie, Phil. Mag. **8**, 1033 (1930).

where $d_s$ and $q_s$ are the stationary stator axes and $a$ and $b$ are the rotor reference axes moving with the same speed as the rotor. Axis $s$ is the shaft axis along which all mechanical quantities (velocity, etc.) are represented.

(b) Experience has shown that the equations of all industrial machinery come out in the simplest form if first the geometric objects of another machine are set up, whose rotor axes are stationary in space (Fig. 1b). This machine is called the "generalized machine" and actually it contains several sets of windings on both stator and rotor. For this machine the salient pole is assumed to be stationary, although in the most general case it is moving.

The transformation matrix $A_\alpha{}^m$ changing the geometric objects of machine (a) to that of machine (b) is given in Eq. (12). The metric tensor of the generalized machine is by $a_{mn}A_\alpha{}^m A_\beta{}^n$

$a_{\alpha\beta}=$

| $\alpha$ \ $\beta$ | $d_s$ | $d_r$ | $q_r$ | $q_s$ | $s$ |
|---|---|---|---|---|---|
| $d_s$ | $L_{d_s}$ | $M_d$ | 0 | 0 | 0 |
| $d_r$ | $M_d$ | $L_{d_r}$ | 0 | 0 | 0 |
| $q_r$ | 0 | 0 | $L_{q_r}$ | $M_q$ | 0 |
| $q_s$ | 0 | 0 | $M_q$ | $L_{q_s}$ | 0 |
| $s$ | 0 | 0 | 0 | 0 | $L$ |

(37)

Its resistance tensor is

$R_{\alpha\beta}=$

| $\alpha$ \ $\beta$ | $d_s$ | $d_r$ | $q_r$ | $q_s$ | $s$ |
|---|---|---|---|---|---|
| $d_s$ | $R_{d_s}$ | 0 | 0 | 0 | 0 |
| $d_r$ | 0 | $R_{d_r}$ | 0 | 0 | 0 |
| $q_r$ | 0 | 0 | $R_{q_r}$ | 0 | 0 |
| $q_s$ | 0 | 0 | 0 | $R_{q_s}$ | 0 |
| $s$ | 0 | 0 | 0 | 0 | $R$ |

(38)

All the components of its Christoffel symbol $[\alpha\beta, \gamma]$ are zero. (If the salient pole rotates, the components of $[\alpha\beta, \gamma]$ are not zero.) Since

$A_\gamma{}^k A_s{}^\delta \frac{\partial A_k{}^\delta}{\partial x^\sigma} =$

| $\gamma$ \ $\delta$ | $d_s$ | $d_r$ | $q_r$ | $q_s$ | $s$ |
|---|---|---|---|---|---|
| $d_s$ | 0 | 0 | 0 | 0 | 0 |
| $d_r$ | 0 | 0 | 1 | 0 | 0 |
| $q_r$ | 0 | −1 | 0 | 0 | 0 |
| $q_s$ | 0 | 0 | 0 | 0 | 0 |
| $s$ | 0 | 0 | 0 | 0 | 0 |

the components of the tensor $P_{\alpha\beta\gamma}$ are (with $s$ as a fixed index)

$P_{s\beta\gamma}=-P_{\gamma\beta s}=$

| $\gamma$ \ $\beta$ | $d_s$ | $d_r$ | $q_r$ | $q_s$ | $s$ |
|---|---|---|---|---|---|
| $d_s$ | 0 | 0 | 0 | 0 | 0 |
| $d_r$ | 0 | 0 | $L_{q_r}$ | $M_q$ | 0 |
| $q_r$ | $-M_d$ | $-L_{d_r}$ | 0 | 0 | 0 |
| $q_s$ | 0 | 0 | 0 | 0 | 0 |
| $s$ | 0 | 0 | 0 | 0 | 0 |

(39)

All the other components of $P_{\alpha\beta\gamma}$ are zero. (It is interesting to note that the Rotation-coefficients of Ricci do represent a rotation by changing the rotor flux-linkage wave to a rotor flux-density wave, at right angles to it in space.)

(c) When a particular industrial machine is given (Fig. 2) its matrix of transformation $A_\sigma{}^\alpha$ is read off the diagram of connections and its geometric objects, $a_{\sigma\epsilon}$, $R_{\sigma\epsilon}$, $[\sigma\epsilon, \omega]$ and $P_{\sigma\epsilon\omega}$ are calculated from those of the "generalized machine" given.

When only one set of brushes exist on a rotor winding, the machine is a nonholonomic *dynamical* system.

(d) Calculation shows that in case of rotating electrical machinery:

(1) Assuming first a holonomic machine (Fig. 1a, in which all reference axes move with the same speed as the conductors, but differ in speed from each other), in going over to a new nonholonomic reference frame *there exists only one set of reference frames for which the nonholonomic object* $\Omega_{\alpha\beta,\gamma}$ (contained in $[\alpha\beta, \gamma]$ of Eq. (6)) *can be replaced by a tensor* $P_{\alpha\beta,\gamma}$. In that set of frames all axes move with identical speed. (Fig. 1b in which all reference axes are stationary.)

(2) The matrix of transformation $A_\sigma{}^\alpha$ existing between the frames of this preferred set contain only *constant* components (the reference axes may form any constant angles with each other).

(3) In going over now from this preferred frame of reference to a general nonholonomic frame of reference, *the nonholonomic object* (contained in $[\alpha\beta, \gamma]$ of Eq. (19)) *cannot again be replaced by a tensor*, since certain terms in the nonholonomic form of $[\alpha\beta, \gamma]$ cancel.

*Hence among all the possible sets of reference frames of rotating electrical machinery there are two preferred sets:*

(1) *The reference axes are rigidly connected to the conductors.*
(2) *The reference axes are rigidly connected to each other.*

These and other details will be covered in another publication.

## VI. Advantages of the Equations in Tensorial Form

In analyzing the performance of any given industrial machine with the aid of canonical equations, the engineer may use either the dynamical Eq. (26), or Eq. (29) in terms of tensor concepts. That is, the engineer may use either the holonomic machine (Fig. 1a) or the generalized machine (Fig. 1b) as his *preferred* reference machine.

*Experience has amply shown that the tensor form of the equation of motion is so far superior to the dynamical form for actual engineering calculations, that the use of the tensorial form is an absolute necessity in engineering and not just a convenience.* All successful rotating-machine methods of analysis (equivalent circuits, two-reaction theories, cross-field and revolving-field theories, etc.) have been based upon the tensorial form (Eq. (29)), though this fact is not generally realized.

(1) In using the dynamical Eq. (26) holonomic (moving) reference axes $n$, $k$ must always be assumed as the starting reference axes. That means that for machines with nonholonomic *stationary* axes (which comprises by far the largest number of industrial machines) the geometric objects of a machine with *moving* axes must be assumed first, containing always functions of $\theta$ (for instance Eq. (36)) and then they must be transformed with the aid of a transformation matrix $A_\alpha{}^m$ (for instance Eq. (12)) whose components are also functions of $\theta$.

In other words, the dynamical Eq. (26) compel the engineer to select as his starting reference system one, whose equations happen to assume a most complicated form. The change from this complex form to the final simple form represents extra mental and physical work for the engineer.

On the other hand the tensorial form of the equations allows the engineer to select as his starting (preferred) reference system any one in which the equations assume the simplest form. That reference system happens to be a quasi-holonomic system, that of the "generalized machine" (Fig. 1b) whose geometric objects are of very simple form and are not functions of $\theta$, as shown in the previous section, Eqs. (37), (38) and (39).

The transformation matrix $A_\sigma{}^\alpha$ representing the step from this generalized machine to an actual industrial machine (Fig. 2) (that is to any allowable reference system) is of the simplest form, usually containing only constant components, hence the calculation of the geometric objects of the actual machine is of simple character.

| $A_\sigma{}^\alpha$ | $d_s$ | $f$ | $s$ |
|---|---|---|---|
| $d_s$ | 1 | 0 | 0 |
| $d_r$ | 0 | $\cos\alpha$ | 0 |
| $q_r$ | 0 | $\sin\alpha$ | 0 |
| $q_s$ | 0 | 0 | 0 |
| $s$ | 0 | 0 | 1 |

Fig. 2. Repulsion motor and its matrix of transformation.

(2) In using the dynamical Eq. (26) the matrix of transformation $A_\alpha{}^m$ in many cases is a *singular* matrix and the inverse transformation $A_m{}^\alpha$, which is needed in (26), cannot be established. In such cases nonholonomic "*equations of constraints*" must be set up in addition to the holonomic dynamical equations and the calculations become still more involved.

However, in transforming the tensorial Eq. (29) of the generalized machine, the inverse of a singular matrix of transformation is not needed, since the indices of all geometric objects are covariant indices.

(3) The *physical picture* of the phenomena taking place inside the rotating machine, as given by the various geometric objects and terms, depends on the starting reference system. When a holonomic machine is assumed as the starting reference system, the physical interpretation of each term for any other machine is most unusual, artificial and is of no aid to the engineer in his visualization.

On the other hand *the physical picture given by the various terms of the tensorial Eq. (29) is the same that has been used by electrical engineers previously in setting up the equations of a machine from purely physical considerations*, without the use of any canonical equations.

(4) It will be shown later in Section VII that the tensor form of Maxwell's field equations also makes it *necessary* to consider the extra term in Eq. (26) as a *tensor* $P_{\alpha\beta\gamma}\dot{x}^\alpha\dot{x}^\beta$ instead of a non-tensor term $2\Omega_{\alpha\beta,\ \gamma}\dot{x}^\alpha\dot{x}^\beta$, in order to correlate the

dynamical equations of Lagrange with the field equations of Maxwell.

There are still other considerations, all pointing to the advantage of the tensorial formulation of engineering problems.

## VII. The Field Equations of Maxwell

(a) If a cross section is made of any rotating machine, it is found that the current-density and flux-density waves are sinusoidally distributed in *space* (neglecting higher space harmonics). This suggests trying to set up Maxwell's field equations in tensor form for the electromagnetic part of the quasi-holonomic system under consideration. It is of course known that Maxwell's equations do not incorporate the effect of mechanical energies and are insufficient to determine the performance of electrical machinery during acceleration.

Calculation shows that the field equations of Maxwell for a quasi-holonomic system can be set up *only* when the tensorial equation of motion of Lagrange is used, Eqs. (19) or (29) which is derived by assuming a *special nonholonomic reference frame* as a starting point, in which all axes have the same velocity.

It is found that in order to get correct answers, both connections $\Gamma_{\alpha\beta,\,\gamma}$ and $\Lambda_{\alpha\beta,\,\gamma}$ have to be considered. Covariant differentiation with respect to $\Gamma_{\alpha\beta,\,\gamma}$ will be denoted by $\overset{G}{\nabla}$ ($G$ = geometrical connection) and that with respect to $\Lambda_{\alpha\beta,\,\gamma}$ will be denoted by $\overset{D}{\nabla}$ ($D$ = dynamical connection). Since only the electromagnetic part of the system is considered, the moment of inertia of the rotor $L$ is replaced by unity.

It should be noted that since the electrostatic field existing in the dynamical system is not considered, the time axis does not appear in the analysis.

(b) The flux-linkage wave (electrokinetic momentum) of all axes is defined by Maxwell[9] as the electromagnetic vector-potential. However it is necessary to assume, that this definition is valid *only* for the generalized machine with stationary axes. That is

$$\varphi_\alpha \overset{*}{=} a_{\alpha\beta}\dot{x}^\beta. \tag{40}$$

[9] Maxwell, *Electricity and Magnetism*, Vol. 2, p. 247.

For the non-Riemannian generalized machine (Fig. 1b).

$$\varphi_\alpha = \begin{array}{c|c|c|c|c} d_s & d_r & q_r & q_s & s \\ \hline \varphi_{d_s} & \varphi_{d_r} & \varphi_{q_r} & \varphi_{q_s} & \dot{x}_s \end{array} \tag{41}$$

where $\varphi_{d_r} = i^{d_s} M_d + i^{d_r} L_{d_r}$, also $\varphi_{q_r} = i^{q_s} M_q + i^{q_r} L_{q_r}$ etc.

For the Riemannian machine (Fig. 1a) $\varphi_m$ *must* be defined in terms of $\varphi_\alpha$, instead of $a_{mn}i^n$

$$\varphi_m = \varphi_\alpha A_m{}^\alpha \tag{42}$$

where $A_\alpha{}^m$ is given in Eq. (12).

$$\varphi_m = \begin{array}{c|c|c|c|c} d_s & a & b & q_s & s \\ \hline \varphi_{d_s} & \begin{array}{l}\varphi_{d_r}\cos x^a + \\ \varphi_{q_r}\sin x^a\end{array} & \begin{array}{l}-\varphi_{d_r}\sin x^a \\ +\varphi_{q_r}\cos x^a\end{array} & \varphi_{q_s} & \dot{x}_s \end{array} \tag{43}$$

(c) The first equation of Maxwell is

$$F_{\alpha\beta} = 2\overset{G}{\nabla}_{[\alpha}\varphi_{\beta]}. \tag{44}$$

Since the curl of a vector is

$$\overset{G}{\nabla}_{[\alpha}\varphi_{\beta]} = \overset{h}{\partial}_{[\alpha}\varphi_{\beta]} - S_{\alpha\beta}^{\cdot\cdot\gamma}\varphi_\gamma, \tag{45}$$

the skew-symmetric electromagnetic tensor is[10]

$$F_{\alpha\beta} = \partial\varphi_\alpha/\partial x^\beta - \partial\varphi_\beta/\partial x^\alpha - 2S_{\alpha\beta}^{\cdot\cdot\gamma}\varphi_\gamma. \tag{46}$$

For the non-Riemannian generalized machine (Fig. 1b), the components of $\varphi_\alpha$ being constants,

$$F_{\alpha\beta} = -2S_{\alpha\beta}^{\cdot\cdot\gamma}\varphi_\gamma = P_{\beta\gamma\alpha}\dot{x}^\gamma. \tag{47}$$

Calculation shows that the components of the skew-symmetric *field* tensor are

$$F_{\alpha\beta} = \begin{array}{c|c|c|c|c|c} \alpha \backslash \beta & d_s & d_r & q_r & q_s & s \\ \hline d_s & 0 & 0 & 0 & 0 & 0 \\ d_r & 0 & 0 & 0 & 0 & \varphi_{q_r} \\ q_r & 0 & 0 & 0 & 0 & -\varphi_{d_r} \\ q_s & 0 & 0 & 0 & 0 & 0 \\ s & 0 & -\varphi_{q_r} & \varphi_{d_r} & 0 & 0 \end{array} \tag{48}$$

representing (twice) the resultant *flux-density* wave of the rotor, at right angles in space to the flux-linkage wave.

[10] Similar formula for $F_{\alpha\beta}$ is derived from a variational principle by Novobatzky, Zeits. f. Physik **89**, 379 (1934), Eq. (25).

It should be noted that if in Eq. (45) the tensor $-S_{\alpha\beta}^{\cdot\cdot\gamma}$ had been replaced by its original definition $\Omega_{\alpha\beta}{}^{\gamma}$ (Eq. (13)), a nontensor object, then the right-hand side of Eq. (47) could *not* have been defined as $F_{\alpha\beta}$, a *tensor*. Hence, *the replacement of* $\Omega_{\alpha\beta}{}^{\gamma}$ by $-S_{\alpha\beta}^{\cdot\cdot\gamma}$ *is required by Maxwell's equation.*

For the Riemannian machine (Fig. 1a) $S_{\alpha\beta}^{\cdot\cdot\gamma}$ is zero and

$$F_{mn}=\partial\varphi_m/\partial x^n-\partial\varphi_n/\partial x^m. \tag{49}$$

By differentiating Eq. (43)

$F_{mn}=$

| $m$ \ $n$ | $d_s$ | $a$ | $b$ | $q_s$ | $s$ |
|---|---|---|---|---|---|
| $d_s$ | 0 | 0 | 0 | 0 | 0 |
| $a$ | 0 | 0 | 0 | 0 | $\varphi_b$ |
| $b$ | 0 | 0 | 0 | 0 | $-\varphi_a$ |
| $q_s$ | 0 | 0 | 0 | 0 | 0 |
| $s$ | 0 | $-\varphi_b$ | $\varphi_a$ | 0 | 0 |

(50)

where

$$\begin{aligned}\varphi_a&=\varphi_{d_r}\cos x^s+\varphi_{q_r}\sin x^s\\ \varphi_b&=-\varphi_{d_r}\sin x^s+\varphi_{q_r}\cos x^s\end{aligned} \tag{51}$$

giving the rotor flux-density wave. $F_{mn}$ could have also been found from $F_{\alpha\beta}$ (Eq. (48)) by a tensor transformation $F_{mn}=F_{\alpha\beta}A_m{}^{\alpha}A_n{}^{\beta}$, where $A_{\alpha}{}^{m}$ is given in Eq. (12).

(d) If Eq. (47) is substituted into the dynamical equation of motion (29), it becomes for *any* machine or quasi-holonomic system

$$a_{\gamma\alpha}\ddot{x}^{\alpha}+[\alpha\beta,\gamma]\dot{x}^{\alpha}\dot{x}^{\beta}+F_{\gamma\alpha}\dot{x}^{\alpha}=f_{\gamma}. \tag{52}$$

This form of the equation of motion of a charged particle is given (without $f_\gamma$) in many papers on the unified field theory.[11] It should be noted that in the *Riemannian* equation of motion (1) there is no place left for the electromagnetic-field quantities.

In case of rotating machinery the term $F_{\gamma\alpha}\dot{x}^{\alpha}$ represents:

(1) The total torque exerted on the rotor ($\gamma=s$)
(2) The rotor generated voltages due to rotation of the conductors.

[11] See the paper by Mandel, reference 7, Eqs. (7) and (10) or Tolman, *Relativity, Thermodynamics and Cosmology* (1935), Eq. (103.1), etc.

The term $[\alpha\beta,\gamma]\dot{x}^{\alpha}\dot{x}^{\beta}$ gives the voltages generated in all stator and rotor windings due to the motion of the reference axes and of the magnetic paths. It also gives the torque on the latter.

(e) From Eq. (44) it automatically follows that

$$\overset{G}{\nabla}_{[\alpha}F_{\beta\gamma]}=0. \tag{53}$$

(f) The *second independent equation of Maxwell* is

$$\overset{D}{\nabla}^{\alpha}F_{\alpha\beta}=s_{\beta} \tag{54}$$

$$s_{\beta}=\partial F^{\alpha}\beta/\partial x^{\alpha}+\Lambda_{\beta\gamma,\ s}F^{\gamma s}. \tag{55}$$

For the non-Riemannian generalized machine

$$s_{\beta}=\Lambda_{\beta\gamma,\ s}F^{\gamma s}=P_{\beta\gamma s}F^{\gamma s}. \tag{56}$$

Calculation shows that

$s_\beta=$

| $d_s$ | $d_r$ | $q_r$ | $q_s$ | $s$ |
|---|---|---|---|---|
| 0 | $\varphi_{d_r}$ | $\varphi_{q_r}$ | 0 | 0 |

(57)

representing the covariant *rotor currents* in the stationary rotor axes. (See Eq. (40).)

For the Riemannian machine the last term of Eq. (55) is zero and

$$s_m=\partial F^{n}{}_{m}/\partial x^{n}. \tag{58}$$

Differentiating Eqs. (50) and (51)

$s_m=$

| $d_s$ | $a$ | $b$ | $q_s$ | $s$ |
|---|---|---|---|---|
| 0 | $\varphi_a$ | $\varphi_b$ | 0 | 0 |

(59)

representing the covariant *rotor currents* in the moving rotor axes.

(g) From Eq. (52) follows the force equation (torque on the rotor)

$$F_{\alpha\beta}s^{\beta}=p_{\alpha}. \tag{60}$$

(h) From Eq. (54) it automatically follows that

$$\overset{D}{\nabla}^{\alpha}s_{\alpha}=0. \tag{61}$$

(i) It should be noted that Maxwell's equations do not define geometric objects for the *stator* axes and $F_{\alpha\beta}$ and $s_\alpha$ have no components along $d_s$ and $q_s$. That is, the *free* indices in Maxwell's equations do not assume stator values (but the dummy indices do).

(j) The *stress-energy tensor* is defined as

$$T_{\alpha\beta}=F^{\dot{\alpha}\epsilon}F_{\beta\epsilon}-\tfrac{1}{4}g_{\alpha\beta}F_{\epsilon\sigma}F^{\epsilon\sigma}. \tag{62}$$

Defining $g_{\alpha\beta}=\sum_m A_\alpha{}^m A_\beta{}^m$ for the rotor axes only, (Eq. (12)) it gives the unit matrix $\delta_{\alpha\beta}$.[12]

For the non-Riemannian machine calculation gives the symmetrical tensor

$$T_{\alpha\beta}=\begin{array}{c|ccccc} & d_s & d_r & q_r & q_s & s \\ \hline d_s & 0 & 0 & 0 & 0 & 0 \\ d_r & 0 & \varphi_{q_r}{}^2-E & -\varphi_{d_r}\varphi_{q_r} & 0 & 0 \\ q_r & 0 & -\varphi_{d_r}\varphi_{q_r} & \varphi_{d_r}{}^2-E & 0 & 0 \\ q_s & 0 & 0 & 0 & 0 & 0 \\ s & 0 & 0 & 0 & 0 & E \end{array} \tag{63}$$

where $E$ is the magnetic energy stored in the rotor

$$E=\frac{1}{2}\left(\frac{\varphi_{q_r}{}^2}{L_{d_r}{}'}+\frac{\varphi_{d_r}{}^2}{L_{q_r}{}'}\right) \tag{64}$$

and $L_{d_r}{}'$ and $L_{q_r}{}'$ are the inductances of the rotor windings measured with the stator windings *short-circuited*

$$L_r{}'=(L_sL_r-M^2)/L_s. \tag{65}$$

The stress-energy tensor represents:

(1) The Maxwell stresses existing between the rotor windings.
(2) The magnetic energy stored in the rotor.

(k) The existence of two (and only two) sets of *preferred* reference systems brings into sharp contrast the two schools of thought that exist in analyzing rotating machine performance: namely those using flux linkages and those using flux densities:

(1) In starting with reference axes moving with the conductors, the machine analysis becomes a problem in *particle* dynamics, using the holonomic equation of motion of Lagrange, Eq. (1) as a starting point. In the analysis only quantities measured at the machine terminals occur, the flux-density wave, the current-density wave and in general the field equations of Maxwell are nonexistent.
(2) In starting with stationary reference axes, moving sinusoidal *waves* inside the machine appear, that were nonexistent in the previous analysis. The quasi-holonomic equations of motion, Eqs. (52) or (19) are used as a starting point and the field equations of Maxwell apply to the space waves.

Hence the *dual point of view*, that is, consideration of a rotating electrical machine as a collection of *particles* (terminal quantities) or as a collection of *waves* (sinusoidal space quantities), is a consequence of the existence of *two preferred reference frames* from which the phenomenon may be viewed.

The *holonomic* preferred frame must be used as a starting point for the dynamical equations of Lagrange, the *quasi-holonomic* preferred frame must be used as a starting point for the field equations of Maxwell.

[12] Levi-Civita, *A Simplified Presentation of Einstein's Unified Field Equations* (Blackie & Son, 1929), Eq. (35).

# The Thermal Conductivity of Tungsten

WALTER C. MICHELS AND MARTHA COX, *Department of Physics, Bryn Mawr College*

(Received December 28, 1935)

The theory of the heat losses from an electrically heated wire, as given by Roberts and Kannuluik, has been extended to include a wire supported by springs. The results obtained have been used in a measurement of the thermal conductivity of commercial tungsten wire in the temperature range from 78°K to 273°K. There are definite indications that the Wiedemann-Franz law breaks down badly in this region.

IN the course of an investigation of accommodation coefficients, recently carried out in this laboratory,[1] it became necessary to use the thermal conductivity of the tungsten wires in order to correct for the end losses in the Pirani gauge used. A search of the literature revealed that, although tungsten has been very thoroughly studied at high temperatures,[2] very little work has been done on its thermal properties in the neighborhood of room temperature. Kannuluik[3] recently reported the value of the thermal conductivity at 273°K as 1.67 watts/cm deg., checking quite well with the older values, and Roberts,[4] in the course of his work on accommodation coefficients, estimated the value at 80°K to be 2.6 watts/cm deg. No determinations were found at other temperatures in this neighborhood. Roberts' value, moreover, failed to check with a rough determination made in the course of our accommodation coefficient work. It became necessary, therefore, to make a set of determinations of the conductivity, and we chose the measurement of the losses from an electrically heated wire in vacuum as the most convenient and reliable method. The values which are reported in this article are essentially preliminary, and it is planned to carry out further determinations over a wider temperature range and with improved apparatus, but we feel justified in presenting them because of the paucity of existing data.

## I. THEORY OF AN ELECTRICALLY HEATED WIRE WITH END SUPPORTS

Kannuluik[3] and Roberts[4] have developed the theory of an electrically heated wire with its ends and surroundings at a common temperature, $T$. In practice, however, the heat treatment of the wire requires that it be supported by springs,[5] and the temperature drop in these introduces an appreciable error in some cases.

Consider a uniform cylindrical wire of radius $a$, length $l$ and resistivity $\rho$, mounted in vacuum, the surroundings being at a temperature $T$. Suppose that the wire is supported by two springs made of the same material, each of length $l_s$ and radius $a_s$. Finally, let the current flowing in the wire be $i$ and the temperature drop in the springs (which is to be evaluated later) be $\Delta T$. In a section of the wire of length $dx$, the electrical power input is $i^2\rho dx/\pi a^2$. If the temperature excess above the ends is $t$, the net radiation loss is

$$2\pi a\epsilon\sigma\{(T+t+\Delta T)^4-T^4\}dx,$$

where $\epsilon$ is the emissivity of the wire and $\sigma$ is the Stefan-Boltzmann constant. When $(t+\Delta T)/T$ is small, this reduces to

$$8\pi a\epsilon\sigma T^3(t+\Delta T)dx.$$

Finally, the conduction loss from the section is equal to

$$-\pi a^2K(d^2t/dx^2)dx,$$

where $K$ is the thermal conductivity of the wire. Equating the power input to the sum of these two losses, we have:

$$d^2t/dx^2-At+B=0, \qquad (1)$$

where $A=8\epsilon\sigma T^3/aK$ and $B=i^2\rho/\pi^2a^4K-A\Delta T$.

Solving this equation, subject to the boundary conditions that $t=0$ at $x=0$ and at $x=1$, we obtain the temperature distribution of the wire:

$$t=\frac{B}{A}\left\{1-\frac{\sinh xA^{\frac{1}{2}}+\sinh(l-x)A^{\frac{1}{2}}}{\sinh lA^{\frac{1}{2}}}\right\}. \qquad (2)$$

[1] Michels and White, Phys. Rev. **47**, 197A (1935).
[2] E.g., see Forsythe and Worthing, Astrophys. J. **61**, 146 (1925).
[3] Kannuluik, Proc. Roy. Soc. **A131**, 320 (1931).
[4] Roberts, Proc. Roy. Soc. **A135**, 192 (1932).
[5] Blodgett and Langmuir, Rev. Sci. Inst. **5**, 321 (1934).

This expression for $t$ enables us to find the total conduction and radiation loss from the wire. First, setting the conduction loss equal to

$$2\pi a^2 K(dt/dx)_{x=0},$$

we find:

$$Q_k = 2\pi a^2 K B f(A)/A^{\frac{1}{2}}, \tag{3}$$

where $f(A)$ has the value $(\cosh lA^{\frac{1}{2}}-1)/\sinh lA^{\frac{1}{2}}$. Finally, the total radiation loss is:

$$Q_r = 8\pi a\epsilon\sigma T^3(\bar{t}+\Delta T)l, \tag{4}$$

where $\bar{t}$ is the average value of $t$, obtained by integrating Eq. (2) over the length of the wire.

It is interesting to note that Eq. (2) gives a parabolic temperature distribution if the radiation or lateral losses are negligible, and a rectangular distribution if the conduction losses are negligible. For the former the temperature gradient at the ends has a magnitude $Bl/2$, while for the latter it is infinite. In the case of any actual wire the value of this gradient will be between these extremes, and will increase as the lateral losses are increased. Consequently, if the lateral losses are made greater by any means, such as increasing the emissivity or surrounding the wire with gas, *the conduction losses also increase for a given average temperature rise.* This is an extremely important factor in determining the accuracy of any measurement involving the heat losses from a wire.

We shall now consider the springs themselves, and shall use the subscript $s$ to denote quantities referring to the springs, otherwise keeping our earlier notation. If we let

$$A_s = 8\epsilon_s\sigma T^3/a_s K, \quad \text{and} \quad B_s = i^2\rho_s/\pi^2 a_s^4 K,$$

Eq. (1) represents the springs. The boundary conditions are now $t_s = 0$ when $x = 0$, and

$$Q_k/2 = \pi a_s^2 K dt_s/dx \quad \text{at} \quad x = l_s.$$

Imposing these conditions, we find that:

$$\Delta T = (t_s)_{x=l_s} = \frac{(B_s/A_s)(\cosh l_s A_s^{\frac{1}{2}}-1) - Q_k/2\pi a_s^2 K A_s^{\frac{1}{2}} \sinh l_s A_s^{\frac{1}{2}}}{\cosh l_s A_s^{\frac{1}{2}}} \tag{5}$$

and

$$\bar{t}_s = \frac{(B_s l_s/A_s)\sinh l_s A_s^{\frac{1}{2}} - (B_s/A_s^{\frac{1}{2}})\cosh l_s A_s^{\frac{1}{2}} + (Q_k/2\pi a_s K_s A_s^{\frac{1}{2}})(\cosh l_s A_s^{\frac{1}{2}}-1)}{l_s \cosh l_s A_s^{\frac{1}{2}}}. \tag{6}$$

In case either the emissivity or the temperature is so low that the radiation losses may be neglected, these equations become simplified. In this case $f(A)/A^{\frac{1}{2}}$ approaches the value $l/2$ and we have:

$$t = (B/2)(lx - x^2), \tag{7}$$

$$\bar{t} = Bl^2/2, \tag{8}$$

$$\Delta T = B_s l_s^2/2 + Q_k l_s/2\pi a_s^2 K, \quad \text{and} \tag{9}$$

$$\bar{t}_s = B_s l_s/4 + Q_k l_s/4\pi a_s^2 K. \tag{10}$$

## II. Determination of the Thermal Conductivity

The gauge used for these measurements consisted of a tungsten wire of 0.00254 cm radius and 14.8 cm long mounted in a Pyrex tube, between springs of 0.00635 cm radius wire. Each spring contained 6.0 cm of wire. The tube was so constructed that it could be lowered into a one liter Dewar flask for thermostatic purposes. After a thorough baking on the pumps (mercury pumps with $CO_2$ and acetone as a freezing agent), the wire was aged by flashing the filament at white heat for several hours. This flashing removes the tungsten oxide from the surface and etches the surface to a stable condition.[1, 6] The tube was then kept on the pumps at a measured pressure of $<10^{-6}$ mm during the measurements.

A standard resistance of one ohm was connected in series with the tube and the potential differences across this resistance and across the tube itself were measured by a potentiometer. A series of currents were used, the maximum giving an increase of resistance of about 25 percent. At two of the three temperatures used, those obtained by thermostating with melting ice and a

[6] Taylor and Langmuir, Phys. Rev. **44**, 423 (1933).

carbon dioxide snow-acetone mixture, the resistance of the tube was found to be a linear function of the power input, but the departure from linearity at the third temperature (that of liquid nitrogen) was distinctly outside the experimental error. Since the resistance is very closely a linear function of temperature and the power input is the sum of $Q_k$ and $Q_r$, we should expect from Eqs. (3) and (4) that this linear relation would hold. Two approximations, however, have been made in deriving these equations, when it was assumed that the resistivity is a constant over the length of the wire and that the radiation loss is proportional to the temperature excess above the surroundings. These two errors compensate somewhat, since the increase in resistivity accounts for a higher power input to those sections of the filament from which the radiation loss is greater than we have assumed. But at the temperature of liquid nitrogen the radiation loss is negligible, and the departure from linearity is undoubtedly due to variation of resistivity with temperature.

From the measured resistance and power input the thermal conductivity was found by successive approximations. Assuming a value of the thermal conductivity and taking the emissivity of the tungsten oxide on the springs as ten times that of clean tungsten, a consistent solution of Eq. (6) was found. From this the resistance of the springs could be determined, so that the resistance and the average temperature rise of the filament itself were obtainable. Hence $Q_r$ could be calculated from Eq. (4) and subtracted from the power input to the filament to give $Q_k$. Putting this quantity into Eq. (5), $\Delta T$ was found. Putting this into Eq. (3), $f(A)/A^{\frac{1}{3}}$ was obtained and the value of $K$ was found graphically from a plot against this quantity. This method was repeated until the assumed value of $K$ and the final value agreed. Values of the emissivity were taken from the data of Langmuir and Taylor[7] and the resistivity from those of Holburn,[8] with which our own measurements agreed after a small correction was made for the resistances of the joints. At the lowest temperature Eqs. (8)–(10) were used for the calculation.

TABLE I.

| TEMPERATURE (deg. K) | THERMAL CONDUCTIVITY (watts/cm deg.) | $K/\sigma T$ (e.s.u.$^2$) |
|---|---|---|
| 78 | 2.26 | $2.36\times10^{-13}$ |
| 194 | 1.60 | 2.98 |
| 273 | 1.88 | 3.87 |

The values of the thermal conductivity found in this way are given in Table I. Because of the large influence of the diameter of the wire on the value of $K$, we estimate the uncertainty in these values to be about ten percent. In the third column of the table, the values of the Wiedemann-Franz "constant" are tabulated. In connection with this quantity it is interesting to note that the value which we obtain for it checks with that obtained by Kannaluik much more closely than do the values of the thermal conductivities obtained in the two experiments, indicating that its value is not seriously affected by impurities which, apparently, change the electrical and thermal conductivities in the same ratio, as would be expected from theoretical considerations. A very serious departure from the ordinary theory of the Wiedemann-Franz law is indicated, however, in that the ratio $K/\sigma T$ seems to be approaching a value of zero as the temperature approaches zero.

Further work on the thermal conductivities of the metals in this temperature range is planned with two objects in view, to obtain more precise values for practical application and to test this interesting and peculiar behavior of the Wiedemann-Franz constant. We expect to have available for this work a sample of extremely pure tungsten wire.

The authors wish to express their gratitude for a gift of the late Marion E. Reilly which made possible the purchase of some of the apparatus used in this work.

[7] Langmuir and Taylor, J. Opt. Soc. Am. **25**, 321 (1935).

[8] Holburn, *International Critical Tables*, Vol. 6, p. 129.

# The Thermal Stress in a Strip Due to Variation of Temperature Along the Length and Through the Thickness

J. N. GOODIER, *Ontario Reseach Foundation, Toronto*

(Received February 6, 1936)

The thermal stress developed by unequal heating of a flat strip is investigated, the temperature varying along the length and through the thickness, but not across the width. The problem is reduced to one of edge tractions only, for which general solutions are already known. From the properties of these solutions it is shown that for heating local to a narrow transverse band across the width, as might occur in welding processes, there is a concentration of stress near the edge of the strip, at the hottest part, and a maximum tensile stress of magnitude $E\alpha T_{\max.}$, when the temperature is constant through the thickness. The corresponding quantity is found also for temperatures varying through the thickness. For steel $E\alpha T_{\max.} \doteqdot 350 T_{\max.}$ pounds per in.$^2$, the temperature being measured in centigrade degrees above that of the cool parts of the strip, and this approaches the yield point of mild steel for a maximum temperature (rise) of the order of 150°C. For a given maximum temperature, the maximum stress is reduced by about 40 percent by keeping one face cold.

## 1.

WHEN a solid is unevenly heated, the tendency of the hotter parts to expand is restrained by the rigidity of the cooler parts, and a state of thermal stress is developed.

The particular case of the thin flat strip heated locally all across the width, is suggested by welding processes. Two types of temperature distribution will be considered.

(i) The temperature, $T$, is constant through the thickness, and across the width. With axes as in Fig. 1, it is a function of $x$ only.

(ii) The temperature on either face ($z=\pm\frac{1}{2}h$) is a function of $x$ only, and the variation through the thickness $h$ from $T_1(x)$ to $T_2(x)$ is linear. The temperature at any point can then be expressed as

$$T=\tfrac{1}{2}(T_1+T_2)+(z/h)(T_1-T_2).$$

This type of distribution will occur when one face of the strip only is heated.

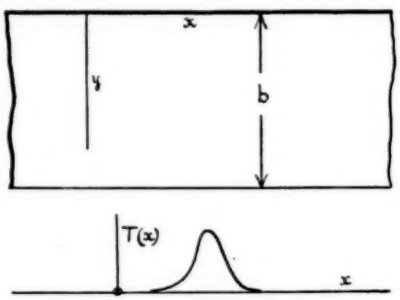


FIG. 1.

## 2. FIRST TYPE OF TEMPERATURE DISTRIBUTION

The distribution (i) will give rise to a state of generalized plane stress. Let $u$, $v$ be the displacements, $X_x$, $Y_y$, $X_y$ stress components, averaged across the thickness. The stress-strain relations are

$$X_x=\frac{E}{1-\sigma^2}\left[\frac{\partial u}{\partial x}+\sigma\frac{\partial v}{\partial y}-(1+\sigma)\alpha T\right],$$

$$Y_y=\frac{E}{1-\sigma^2}\left[\frac{\partial v}{\partial y}+\sigma\frac{\partial u}{\partial x}-(1+\sigma)\alpha T\right],$$

$$X_y=\frac{E}{2(1+\sigma)}\left(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\right),$$

where $E$ is Young's modulus, $\sigma$ Poisson's ratio, and $\alpha$ the coefficient of linear expansion. When these are substituted in the equations of equilibrium, the result is the two equations

$$\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y}\right)\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)+\frac{1-\sigma}{1+\sigma}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)(u,v)$$
$$=2\alpha\left(\frac{\partial T}{\partial x},\frac{\partial T}{\partial y}\right).$$

When $T$ is a function of $x$ only, we find a particular integral by assuming that $u$ is a function of $x$ only, and $v=0$. This leads to $\partial u/\partial x=(1+\sigma)\alpha T$, and to the stress components

$$X_x=X_y=0, \quad Y_y=-E\alpha T. \tag{1}$$

There is therefore normal traction $-E\alpha T$ on the edges $y=0$, $b$. In order to make these edges free

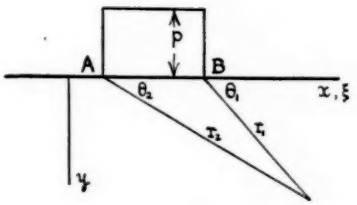


FIG. 2.

it is necessary to superpose on (1) the stress distribution due to equal and opposite edge tractions ($Y_y = E\alpha T$ on $y=0, b$) and the resultant distribution will then be the thermal stress in the strip.

The problem of the strip of finite width loaded by edge forces has been completely solved,[1] and formulas for the thermal stress can now be written down.

3.

We shall consider in detail only the case when the region of high temperature is narrow, in the $x$ direction, in comparison with the width of the strip. The traction $E\alpha T$ is then concentrated on a small length of the edge, and the stress it produces will be local to this length. The existence of the other edge is immaterial, and each half of the strip can be treated as a half-plane. The stress in the half-plane due to normal traction $p$ constant over a segment $AB$ of the edge $y=0$ (Fig. 2) is derivable[2] from the Airy stress function $\chi = (p/2\pi)(r_1^2\theta_1 - r_2^2\theta_2)$ by means of the formulas

$$X_x = \partial^2\chi/\partial y^2, \quad Y_y = \partial^2\chi/\partial x^2, \quad X_y = -\partial^2\chi/\partial x \partial y.$$

Then

$$\begin{aligned} X_x &= (p/2\pi)[2(\theta_1-\theta_2)+\sin 2\theta_1 - \sin 2\theta_2], \\ Y_y &= (p/2\pi)[2(\theta_1-\theta_2)-\sin 2\theta_1 + \sin 2\theta_2], \\ X_y &= (p/2\pi)[\cos 2\theta_1 - \cos 2\theta_2]. \end{aligned} \tag{2}$$

It is clear that these satisfy the boundary condition: $Y_y = p$ within $AB$, zero outside $AB$; and $X_y = 0$.

Making the segment infinitesimal, and introducing $\xi$ as the coordinate of $A$ measured along the edge, the stress function becomes $(p/2\pi)\cdot(\partial/\partial\xi)r'^2\theta'\cdot d\xi$ where $r'^2 = (x-\xi)^2+y^2$, and $\theta'$ is the principal value of $\tan^{-1}(y/x-\xi)$. Carrying out the differentiation, and integrating the effects of all the elements of a load function $p(\xi)$ we find that the corresponding stress function is

$$\chi(x, y) = -\frac{1}{\pi}\int_{-\infty}^{+\infty} p(\xi)(x-\xi)\tan^{-1}\frac{y}{x-\xi}d\xi \tag{3}$$

except for a part which makes no contribution to the stress. Traction $E\alpha T$ is obtained by substituting $p(\xi) = E\alpha T(\xi)$.

From (2) we obtain, for the uniformly loaded segment, $X_x = Y_y$ on the edge, where $Y_y$ is of course equal to the applied traction $p$. It follows that at the edge $X_x$ is zero outside the loaded segment; when we assemble different loadings on a series of infinitesimal segments, so as to build up the arbitrary loading $p(\xi)$ or $p(x)$, the stress component $X_x$, at the edge, within each infinitesimal segment is simply due to the load on the segment itself, and we shall have in the resultant stress distribution, defined by (3), $(X_x)_{y=0} = p(x)$.

The normal edge traction $E\alpha T$ therefore produces an edge stress $X_x = E\alpha T$. There is no contribution to $X_x$ from (1), so that this is the actual component of the thermal stress. The heating being local, the stress due to the edge tractions diminishes rapidly from the edge inwards. If the temperature rise is appreciable only between $x = \pm\lambda$ the stress-component $X_x$ will be negligible, in comparison with the edge value, beyond distances from the edge of the order $2\lambda$. In the interior the principal part of the stress will be given by (1).

The character of the components $X_x$ and $Y_y$ across the middle section ($x=0$) is indicated in Fig. 3. The component $X_x$ is tensile at the edges with a maximum value $E\alpha T_{\max.}$. For steel this is about $350T_{\max.}$ pounds per in.$^2$, $T_{\max.}$ being measured in degrees centigrade above the unheated part of the strip. The degree of heating required for a tensile stress of 50,000 lbs. per sq. in.—near the yield point of an ordinary steel, is given by $T_{\max.} = 143°C$. Further heating would lead, in such a material, to permanent deformation, and a residual stress would remain after cooling.

[1] L. N. G. Filon, Phil. Trans. Roy. Soc. **A201**, 63 (1903); and T. von Kármán, Abh. Aerodynam. Inst., Tech. Hochschule, Aachen **7**, 3 (1927).

[2] J. H. Michell, Proc. Lond. Math. Soc. **34**, 139 (1901).

## 4. Second Type of Temperature Distribution

The second type of temperature distribution, (ii) of §1, involves both expansion of the strip in its own plane, and flexure transverse to it. The expansion, associated with the mean temperature $\frac{1}{2}(T_1+T_2)$, produces stress of the type already discussed, and is independent of the flexure. But the flexure, associated with the temperature difference, $T_1-T_2$, will in general be influenced by the mean stress due to the mean temperature.

To form the differential equations of the flexure, let $G_1$, $G_2$ be the bending stress couples per unit length on sections normal to the $x$ and $y$ axes, respectively, $H_1$, the torsional stress couple on a section normal to the $x$ axis, $N_1$, $N_2$ the shearing stress resultants per unit length, and $D$ the flexural rigidity $Eh^3/12(1-\sigma^2)$. The equations of equilibrium are

$$\begin{aligned} &\partial G_1/\partial x-\partial H_1/\partial y-N_1=0,\\ &\partial G_2/\partial y-\partial H_1/\partial x-N_2=0,\\ &\frac{\partial N_1}{\partial x}+\frac{\partial N_2}{\partial y}+X_x\frac{\partial^2 w}{\partial x^2}+2X_y\frac{\partial^2 w}{\partial x\partial y}+Y_y\frac{\partial^2 w}{\partial y^2}=0, \end{aligned} \tag{3}$$

the last expressing the equilibrium of the transverse force components (i.e., in the $z$ direction); $X_x$, $X_y$, and $Y_y$ are the mean stress components due to the mean temperature, and $w$ is the transverse deflection. The relations between the stress couples, stress resultants, and the deflection $w$ are

$$\begin{aligned} G_1&=-D\left[\frac{\partial^2 w}{\partial x^2}+\sigma\frac{\partial^2 w}{\partial y^2}-\frac{\alpha}{h}(1+\sigma)(T_1-T_2)\right],\\ G_2&=-D\left[\frac{\partial^2 w}{\partial y^2}+\sigma\frac{\partial^2 w}{\partial x^2}-\frac{\alpha}{h}(1+\sigma)(T_1-T_2)\right],\\ H_1&=D(1-\sigma)\frac{\partial^2 w}{\partial x\partial y},\\ N_1&=-D\frac{\partial}{\partial x}\left[\nabla^2 w-\frac{\alpha}{h}(1+\sigma)(T_1-T_2)\right],\\ N_2&=-D\frac{\partial}{\partial y}\left[\nabla^2 w-\frac{\alpha}{h}(1+\sigma)(T_1-T_2)\right]. \end{aligned} \tag{4}$$

The formulas for $G_1$, $G_2$, for instance, are derived from the consideration that the unrestrained thermal expansion of an element of the strip would produce spherical curvature $(T_1-T_2)\alpha/h$. The couples are responsible not for the whole actual curvature, but only for the difference between it and the free thermal curvature. Elimination of $G_1$, $G_2$, $H_1$, $N_1$, $N_2$, from (3) and (4) leads to

$$D\cdot\nabla^4 w-X_x\frac{\partial^2 w}{\partial x^2}-2X_y\frac{\partial^2 w}{\partial x\partial y}-Y_y\frac{\partial^2 w}{\partial y^2}=D\frac{\alpha}{h}(1+\sigma)\nabla^2(T_1-T_2). \tag{5}$$

$\nabla^2$ stands for $\partial^2/\partial x^2+\partial^2/\partial y^2$ and $\nabla^4\equiv\nabla^2\nabla^2$.

This is the general differential equation for the flexure of a thin plate due to uneven heating, in the sense that $T_1$, $T_2$ may be functions of both $x$ and $y$, the variation through the plate being linear.

The functions giving the mean stress components $X_x$, $X_y$ and $Y_y$ are determined by the equations of generalized plane thermal stress, and the boundary conditions, independently of the flexure.

In the problem we are considering, the temperatures $T_1$ and $T_2$ are functions of $x$ only. Let the strip be heated and subject to the normal edge traction $-E\alpha\cdot\frac{1}{2}(T_1+T_2)$. Then, by (1), $X_x=X_y=0$ and $Y_y=-E\alpha\cdot\frac{1}{2}(T_1+T_2)$. The Eq. (5) reduces to

$$D\nabla^4 w+E\alpha\cdot\frac{1}{2}(T_1+T_2)\frac{\partial^2 w}{\partial y^2}=D\frac{\alpha}{h}(1+\sigma)\frac{\partial^2}{\partial x^2}(T_1-T_2).$$

A particular integral is obtained by taking $w$ to be

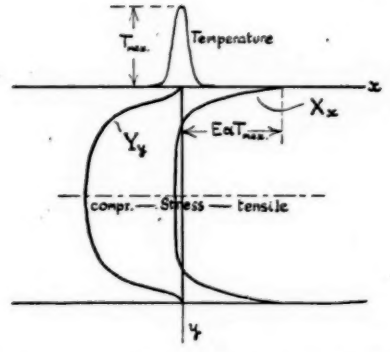


Fig. 3. Character of stress distribution in hottest section of locally heated strip.

a function of $x$ only, in which case the equation is satisfied provided

$$\partial^2 w/\partial x^2 = (\alpha/h)(1+\sigma)(T_1 - T_2).$$

This state of flexure corresponds to

$$G_1 = H_1 = N_1 = N_2 = 0;$$
$$G_2 = (1/12)E\alpha h^2(T_1 - T_2). \quad (6)$$

It is maintained by bending couples $(1/12)E\alpha h^2(T_1 - T_2)$ per unit length and normal traction $-E\alpha \cdot \frac{1}{2}(T_1 + T_2)$, on the edges, together with the given temperature distribution.

To make the edges free we have to superpose equal and opposite edge actions, with the stresses they produce in the interior. The effects of the mean normal traction $E\alpha \cdot \frac{1}{2}(T_1 + T_2)$ can be worked out as in §3, with $T = \frac{1}{2}(T_1 + T_2)$. The effects of edge couples $-(1/12)E\alpha h^2(T_1 - T_2)$ remain to be considered.

The effects of an arbitrary distribution of couples on the strip of finite width could be found by methods analogous to those employed by Filon and von Kármán in the solution of the plane stress problem. We shall consider here only the case when the couples are so localized that one-half of the strip can be treated as a half-plane.

The deflection of the half-plane $y > 0$ due to a concentrated bending couple $M$ at the origin is:[3]

$$w = \frac{M}{D\pi(1-\sigma)(3+\sigma)}[2x \log r - (1+\sigma)y\theta] \quad (7)$$

and the deflection due to a distribution of bending couples, $G$ per unit length, between $A$ and $B$ (Fig. 2), obtained by integrating (7), is

$$w = \frac{G}{D\pi(1-\sigma)(3+\sigma)}[r_1^2 \log r_1 \cdot \sin 2\theta_1 - r_2^2 \log r_2 \cdot \sin 2\theta_2 - 2y^2(\theta_1 - \theta_2) + \tfrac{1}{2}(1+\sigma)(r_1^2\theta_1 - r_2^2\theta_2)]. \quad (8)$$

The deflection due to an arbitrary distribution of edge couples follows from (7) or (8) by integration. On evaluating the couples $G_1$ and $G_2$ corresponding to (8), it is found that at the edge, (where of course $G_2 = G$ within $AB$, and $G_2 = 0$ outside $AB$),

$$G_1 = -(1-\sigma)G/(3+\sigma) \text{ within } AB,$$
$$\text{and} \quad G_1 = 0 \text{ outside } AB.$$

Thus, when we assemble couples on a series of infinitesimal segments of the edge, the couple $G_2$ at the edge, in any segment, is due entirely to the couple acting on the segment itself, and for any distribution of applied couple $G(x)$ we shall have

$$(G_1)_{y=0} = G(x)(1-\sigma)/(3+\sigma).$$

The distribution $G(x) = -(1/12)E\alpha h^2(T_1 - T_2)$ occurring in the thermal stress problem therefore gives rise to the stress couple

$$G_1 = (1/12)E\alpha h^2[1-\sigma)/(3+\sigma)](T_1 - T_2)$$

in the edge. The extreme fiber stress corresponding to this is

$$\pm\tfrac{1}{2}E\alpha[(1-\sigma)/(3+\sigma)](T_1 - T_2).$$

There is no contribution to this stress component from (6). The contribution $(X_x)$ from the mean stress is $E\alpha \cdot \frac{1}{2}(T_1 + T_2)$. The extreme fiber stress in the edge has therefore the values

$$\tfrac{1}{2}E\alpha\{T_1 + T_2 \pm [(1-\sigma)/(3+\sigma)](T_1 - T_2)\}, \quad (9)$$

independent of the thickness. With localized heating the stress due to the temperature difference will have the same kind of distribution in the interior as the mean stress, and can be visualized from Fig. 3, $X_x$ representing $G_1$, and $Y_y$, $G_2$. There will be, with $T_1 > T_2$, a tensile stress in the edge, given by (9) when the positive sign is taken.

As an example to illustrate heating on one face only, let the other face remain cool, so that $T_2 = 0$. Then the maximum tensile stress set up is $2E\alpha T_{1\max.}/(3+\sigma)$. For the same maximum temperature this is, for steel, about 0.61 of the stress $E\alpha T_{\max.}$ set up by heating uniformly through the thickness. It may thus be advantageous to keep one face of the strip as cool as possible.

[3] A. Nadai, *Elastische Platten*, p. 204.

# Variations of the Rigidity Modulus with Changes in Applied Stress

GLADYS M. FRANCIS, *New Jersey College for Women*

(Received December 30, 1935)

An attempt to determine by an interferometer method whether or not any variations occurred in the rigidity modulus similar to those which occurred in Young's modulus in D. K. Froman's experiment. Any variations of more than 0.1 of a fringe could readily be detected. No such variations occurred.

## INTRODUCTION

THE purpose of the present experiment is to discover whether or not any changes occur in the rigidity modulus when stresses are applied in successive steps, starting with a small stress applied in one direction and increasing the stress in the opposite direction until the elastic limit of the material under investigation has been reached. The elasticity is measured after each increase in stress by adding a small additional weight and observing the twist produced in the rod. Any changes in the rigidity modulus taking place would appear in the measurement of the twist.

This investigation was suggested by an experiment performed by Darol K. Froman of the Ryerson Physical Laboratory, University of Chicago.[1] Froman, using an interferometer method for determining the strain, showed that for small stresses, Hooke's law did not hold when a determination of Young's modulus was made, but that Young's modulus had a maximum at a small but finite stress. This maximum occurred when the stress was approximately $3\times10^6$ dynes per sq. cm. Since the constancy of the rigidity modulus is a proof of Hooke's law, the present experiment was performed to discover whether or not the same or similar deviations occurred when a determination of this modulus was made.

## EXPERIMENT

In this experiment, the same general method that Froman followed in his experiment for determining the strain was used; namely, the interferometer method.

The experimental rod $L$ (Fig. 1) was placed inside a brass tube $H$. The end $G$ of the rod $L$ fitted loosely in a hole in a vertical support rod $E$, while the other end $F$ was clamped to a metal piece which slid into the openings of a beam of a laboratory balance. The knife-edge of the balance, which was in a line with the center of the rod, rested in an agate groove which was clamped to the support rod $Q$. Both of the support rods were fastened rigidly to clamps on an optical bench so that the experimental rods could easily be changed without making serious changes in the set-up.

At the ends of the beam of the laboratory balance were suspended the pans $X$ and $Y$. The torque was applied to the rod by adding weights to pan $X$. The end $G$ of the rod $L$ was prevented from turning by means of the clamps $I$, $J$, $K$, and $N$ joined by rods. Each time a different mass was placed in pan $X$ these clamps were adjusted until the beam of the balance was level.

Since the masses added at pan $X$ produced a turning moment at $N$ of the support $E$, another balance pan $A_1$ was added at $P$, the arm $(ef)$ being equal to arm $(cm)$ of the applied force. By adding to this pan $A_1$ at a distance $(ef)$ masses equal to those on pan $X$ a couple resulted so that the system balanced.

At $B$, the tube $H$ was tightly fastened to the rod $L$ by a clamp which was arranged so that it touched the circumference of the rod in a line lying in a plane perpendicular to the axis of the rod.

As shown in the diagram (Fig. 1) the tube $H$ was not as long as the experimental rod so that arrangements for supporting the end $C$ of the tube were necessary. This was done by fastening a brass bar $W$ to the tube at the point $C_1$. To the end $A$ of the bar a knife-edge was fastened.

[1] Froman, Phys. Rev. **35**, 264 (1930).

This knife-edge rested in an agate groove in line with the axis of the rod.

At $C_1$ a brass bar $S$ was attached to the tube $H$. At the end $U$ of the bar, one of the half-silvered mirrors $M$ was fastened. The other half-silvered mirror $M_1$ was suspended by a bar $V$ which was tightly clamped to the experimental rod $L$ at the point $D$. This clamp was constructed so that it touched the rod in the same way that the clamp at $B$ did. The object of these line contacts between the rod and clamps was to furnish a definite beginning and end for measuring the length of the rod. Since the arm of the mirror $M$ was connected to the tube $H$ which was clamped to the rod $L$ at $B$, and since the arm of the mirror $M_1$ was attached to the rod $L$ at $D$, any torque applied to the rod producing a twist between $B$ and $D$ tended to cause the mirrors to separate.

The mirrors were placed with their centers 10.3 cm from the center of the rod $L$ in order to increase the relative change in distance between them when the elasticity of the rod was measured.

In order to prevent a torque from being produced by the weight of the mirror $M$ and the bar $S$, it was necessary to arrange counter weights $O$ and $P_1$, on opposite sides of the tube to balance the system. The counter weight $O$ was fastened at the opposite end of the bar $S$ to which the mirror $M$ was attached and was added to balance the tube on the knife-edges so that the weight of the tube was concentrated at the axis of the rod $L$. The counter weight $P_1$ was fastened to the bar $S$ at the same point as the mirror $M$. This weight was added so as to bring the center of gravity of the mirror and bar $S$ to the axis of the rod.

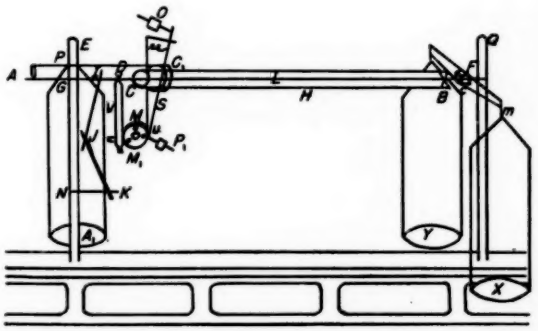


FIG. 1. Arrangement of apparatus to measure torque.

As a source of light a mercury arc lamp, operated on alternating current, was used with filters of didymium and H. R. Yellow to cut out all the mercury lines with the exception of the green one. Since the apparatus in Fig. 1 was very sensitive, the lamp was placed on a board which was suspended by springs in order to damp vibrations which might be transmitted to the apparatus.

Between the lamp and mirrors was placed a large lens to convert the beam of light into a parallel beam. A metal sheet about a foot square with a circular hole of diameter one inch was supported between the lens and the mirrors to reduce the size of the beam of light so that it was smaller than the mirrors.

The fringes, which were circular, were observed through a telescope, the position of which was kept as nearly constant as possible to ensure that the same point on the mirrors was observed each time.

The elasticity of the rod was measured by applying a small additional mass to pan $X$ and observing the fringe shift which it caused. The small mass was suspended by a ring from a hook above the pan of the balance. A support was arranged in such a way that by raising it, the mass would be lifted from the ring, and by lowering it, the mass would be added to the ring.

It was essential that the small additional mass, used for measuring the elasticity, strike the support quite squarely; otherwise, the mass would be set swinging, which in turn set the whole apparatus vibrating, making it impossible to count the number of fringes.

In the arrangement of the present experiment it was possible to obtain readings for stresses applied in either direction since masses could be added to either pan. To find out what happened as the rod was twisted through a point when no torque was applied, a small mass was added to pan $Y$ and the elasticity determined for successive additions of masses to pan $X$.

Seven rods were used in the experiment, of which five, the aluminum, copper, iron, steel and quartz rods, were solid and two, glass and steel, were hollow. These rods were regular commercial rods that are usually sold and were approximately one meter long. The diameters averaged 0.95 cm.

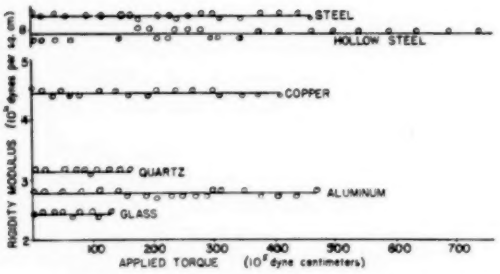


FIG. 2. Graph showing value of rigidity modulus as related to applied torque in $10^5$ dyne cm.

The small additional mass varied for different rods. For the aluminum, copper, hollow steel and quartz rods, it was one gram. For iron and solid steel rods, it was three grams, and for the hollow glass rod it was one-half a gram. The mass used depended upon the number of fringes which were obtained, and was such that the fringe count was approximately twenty-five, so as to make results as accurate as possible. It is estimated that the fringe count is accurate to 0.1 of a fringe.

DATA

| | |
|---|---|
| The distance between clamps in all cases '$L$' | 90.5 cm |
| The wave-length of light used in all cases | 5461 A |
| Arm of applied force | 15.8 cm |
| Arm of mirrors (from center of tube to center of mirror) | 10.3 cm |
| Weight of ring and wire support from which suspended the small additional mass | 0.681 g |

## DISCUSSION

A determination of the absolute value of the rigidity modulus was not of primary interest in this experiment but it was desired to discover whether there were variations in the modulus with changes in the applied stress. It was not necessary in this experiment, as it would have been in the determination of the absolute value of the rigidity modulus, to have the rods uniform or to make exact measurements of the length of the rods and arm of the applied stress. These factors were kept constant so that if any variations had occurred they would have appeared in the fringe count. The slight deviations which were obtained were due to experimental errors in the fringe count since they appeared very irregularly and occurred at different points when measurements on the rods were repeated.

In a determination of the rigidity modulus, the outer layers of the metal rod take more of the shearing stress than the inner layers do. It would be expected, therefore, that when the shearing stress is gradually increased the rigidity modulus for the outer layers would reach a maximum value before the inner layers would, provided such a maximum were to occur. This would mean that decreasing values of the rigidity modulus for the outer layers might be occurring at the same time that the inner layers were increasing to a maximum. The effect of the layers for which the rigidity modulus is increasing would thus counteract the effects of those for which the rigidity modulus is decreasing. Therefore, it would be expected that the value of the rigidity modulus would remain at a maximum value until most of the layers had been subjected to the torque which gives this maximum, but when this had occurred the value of the rigidity modulus would decrease to a constant value. However, for the solid rods this would probably never occur because the elastic limit of the rod would be reached before this would happen.

In view of this then, if any variations were to occur, it might be expected that they would be greater for a tube than for a solid cylinder. It was for this reason that the tubes were used in this experiment.

However, no variations in the rigidity modulus occurred in either the solid rods or tubes. It is therefore concluded that the modulus of rigidity is constant when the stress is gradually increased from a small stress in one direction through a point in the opposite direction where the stress is zero and then up to the elastic limit of the material under investigation.

I wish to express my appreciation to the New Jersey College for Women for the use of the laboratory and to Dr. Frank R. Pratt for his cooperation and helpfulness in this experiment.

# The Synthesis of Gear Tooth Curves*· †

J. STILES BEGGS, *Kodak Research Laboratories, Rochester, New York*

(Received October 24, 1935)

This paper sets forth a method of arriving analytically at a pair of gear tooth curves which will fulfill certain prescribed conditions. The method may also be used to find a curve which will mesh with a given curve and produce a constant angular velocity ratio, e.g., the mating curve for a circular pin tooth.

THE analytical conditions to be met by gear tooth curves may be listed as follows:

I. The primary requirement is that the two gears shall rotate with a constant angular velocity ratio, i.e., the pitch circles[1] shall roll on each other without slipping.

II. Referring to Fig. 1, the point of tangency of the pitch circles, i.e., the pitch point, $M$, is the instantaneous center of either gear with respect to the other.

III. The point of contact (tangency) of the tooth profiles at any instant is denoted by $N$. Since the tooth profiles are in rolling contact and in general, slipping, the line, $MN$, must be normal to both tooth curves at each successive point of contact, by condition II, otherwise, the teeth would either cease to make contact or else tend to pass through each other (Fig. 1).

## MATHEMATICAL FORMULATION OF THE PROBLEM

For the purpose of mathematical analysis, it will be convenient to think of one gear as fixed in the system, $XOY$ (Fig. 2), and the other in the system, $xoy$.

Let $a+b=c$

$$\text{by (1)} \quad a\theta=b\phi; \text{ therefore,} \tag{1}$$

$$\theta+\phi=c/b\cdot\theta. \tag{2}$$

The equations of transformation between the two systems are as follows:

$$x=(Y-c\sin\theta)\sin c/b\theta + (X-c\cos\theta)\cos c/b\theta, \tag{3}$$

$$y=(Y-c\sin\theta)\cos c/b\theta - (X-c\cos\theta)\sin c/b\theta. \tag{4}$$

Let the equation of the tooth curve of the gear in $XOY$ (Fig. 1) be given by

$$Y=F(X). \tag{5}$$

Then, by condition III, the point of contact, $N$, lies at the foot of the perpendicular from $M$ to the curve of Eq. (5).

Obviously, $N$ will trace out in $xoy$ the pinion tooth curve. Let this be given by

$$y=f(x). \tag{6}$$

Eq. (6) may be found by taking the coordinates $(X, Y)$, of $N$ and transforming them by (3) and (4) to the system $xoy$, as parametric equations in $\theta$.

The equation of any straight line through $M$ is given by

$$Y=a\sin\theta+k(X-a\cos\theta). \tag{7}$$

Since it is perpendicular to $Y=F(X)$, at $N$, by

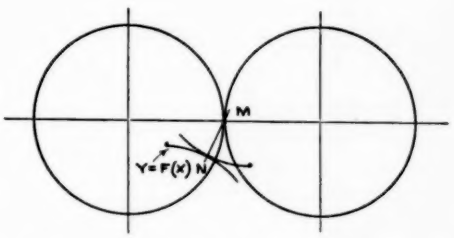


FIG. 1. The general form of gear tooth curve. $M$ is the pitch point and $N$ the point of contact of the curves.

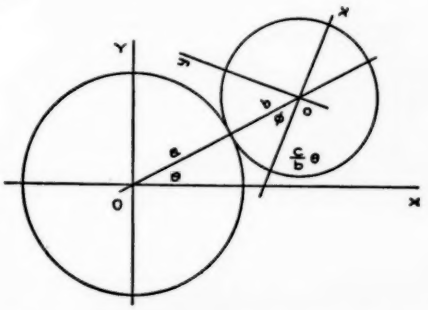


FIG. 2. The gear is considered as fixed in the coordinate system $XOY$ and the pinion in the system $xoy$.

* A revision of a paper read before the Mathematics Club of the University of Alberta.

† Communication No. 560 from the Kodak Research Laboratories.

[1] For the meaning of technical terms, see, e.g., *Machinery's Handbook* (Industrial Press, New York, 1934).

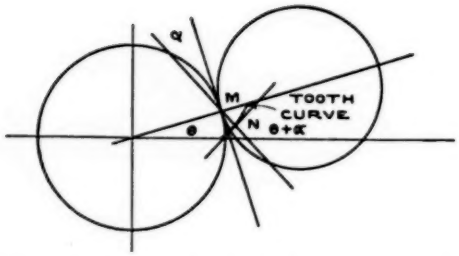


FIG. 3. The form of tooth which gives a constant bearing load at constant torque. This is found to be the involute.

condition III,

$$k=-1/dF(X)/dX, \text{ at the point } N. \quad (8)$$

Therefore, (7) becomes

$$Y=a\sin\theta-1/F'(X)(X-a\cos\theta). \quad (9)$$

The coordinates of $N$ are given by the simultaneous solution of (5) and (9).

## THE INVOLUTE TOOTH

As an example of this method, we impose the following conditions: The reaction on the bearings is to be constant for a constant torque, since a variation in this force would tend to set up vibration and noise. This means that the normal $MN$ must make a constant angle, $\alpha$, for instance, with the line of centers of the gears.

Let Eq. (5) of the required gear tooth curve be given by

$$X=X(\theta), \quad (10)$$

$$Y=Y(\theta). \quad (11)$$

From Fig. 3,

$$dY/dX=\tan(\theta+\alpha). \quad (12)$$

Therefore, (7) becomes

$$Y=a\sin\theta-\frac{X-a\cos\theta}{\tan(\theta+\alpha)}. \quad (13)$$

Then, differentiating (13),

$$dY=a\cos\theta d\theta+(X-a\cos\theta)\times\sec^2(\theta+\alpha)\cot^2(\theta+\alpha)d\theta-\cot(\theta+\alpha)dX-a\sin\theta\cot(\theta+\alpha)d\theta. \quad (14)$$

The simultaneous solution of (12) and (14) gives the differential equation of the locus of $N$. Substituting from (12) in (14) for $dY$, we have

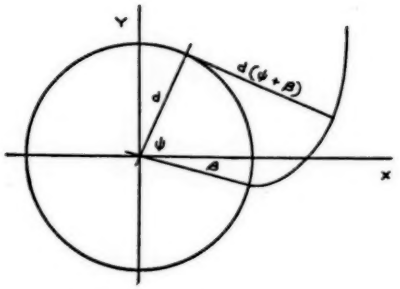


FIG. 4. The involute curve.

$$\frac{dX}{d\theta}-\frac{\cos(\theta+\alpha)}{\sin(\theta+\alpha)}X=-a\cos\alpha\frac{\cos^2(\theta+\alpha)}{\sin(\theta+\alpha)}, \quad (15)$$

which is a linear differential equation whose solution is

$$X=a\cos\alpha[(\theta+\tan\alpha)\sin(\theta+\alpha)+\cos(\theta+\alpha)]. \quad (16)^2$$

Similarly,

$$Y=a\cos\alpha[\sin(\theta+\alpha)-(\theta+\tan\alpha)\cos(\theta+\alpha)]. \quad (17)$$

Referring to Fig. 4, the general equation of the involute is given by

$$X=d\cos\psi+d(\psi+\beta)\sin\psi, \quad (18)$$

$$Y=d\sin\psi-d(\psi+\beta)\cos\psi. \quad (19)$$

Comparing (16) with (18), let

$$a\cos\alpha=d, \quad (20)$$

$$\theta+\alpha=\psi, \quad (21)$$

$$\alpha+\beta=\tan\alpha=(a^2-d^2)^{\frac{1}{2}}/d. \quad (22)$$

The curve (16) and (17) is, therefore, an involute with the base circle, $a\cos\alpha$, and starting at an angle, $\alpha=\tan\alpha-\alpha$, below the $X$ axis, the constants being so chosen as to make the pitch point lie on the $X$ axis.

By imposing certain initial conditions, we have thus arrived at a unique solution of our problem, i.e., the involute is the only curve which will satisfy these conditions.

To find the mating tooth curve, we transform Eqs. (18) and (19) by means of (3) and (4), obtaining,

$$x=-b/a\cdot d[\cos(-a/b\theta+\alpha)+(-a/b\theta+\tan\alpha)\sin(-a/b\theta+\alpha)], \quad (23)$$

[2] I am indebted to Miss Helen Shea of the University of Toronto for the formal integration of Eq. (15).

$$y=-b/a\cdot d[\sin(-a/b\theta+\alpha)-(-a/b\theta+\tan\alpha)\cos(-a/b\theta+\alpha)]. \quad (24)$$

The mating curve is also seen to be an involute.

As slipping is an important factor in the wear of gear teeth, it is desirable to have an expression for it. Referring to Fig. 1, *the slip* is seen to be equal to $MN$ multiplied by the relative angular movement and is, *therefore, zero, only when the point of contact coincides with the pitch point.*

Let $dS$ be an element of arc of the involute gear tooth, and $ds$ be an element of arc of the pinion tooth. If $\sigma$ represents the slip, then

$$d\sigma=dS-ds=d(1+a/b)\theta d\theta. \quad (25)$$

The rate of slipping $$d\sigma/d\theta=d(1+a/b)\theta. \quad (26)$$

Slip from $$\theta=\sigma \text{ to } \theta=\int_{\theta=0}^{\theta=\theta} d\sigma=\frac{d\cdot c}{2b}\theta^2. \quad (27)$$

In practice, to facilitate the cutting of the teeth, standards have been set up. The pressure angle, $\alpha$, is taken as $14\frac{1}{2}°$ and 20°. From Eqs. (18) and (19), we see that the angle of obliquity and pitch diameter determine the shape of the tooth. The diametral pitch determines only its size.

## Mating Curves

This method may also be used to find the mating curve to any given curve, (5), e.g., suppose we wish to find the shape of tooth which will mate with a circular pin tooth and satisfy condition I. In this case, (5) is of the form (Fig. 5):

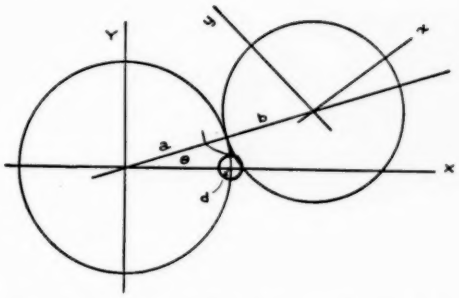


FIG. 5. The pin tooth and its mating curve.

$$X=a-d\sin\theta/2, \quad (28)$$

$$Y=d\cos\theta/2. \quad (29)$$

Transforming by Eqs. (3) and (4), we obtain the mating curve

$$x=a\cos c/b\theta+d\sin(1/2+a/b)\theta-c\cos a/b\theta \quad (30)$$

$$y=-a\sin c/b\theta+d\cos(1/2+a/b)\theta+c\sin a/b\theta, \quad (31)$$

which is not a circle.

It is interesting to note that if we replace the pin by a ball bearing of the same diameter and thus introduce at least three balls and an outer race unfixed in either $XOY$ or $xoy$, we obtain a pair of gears in which there is no sliding friction.

Any form of (5), such as an involute or cycloid, which will undergo the transformation (3) and (4) without change of form will yield a mating tooth of similar shape, or, if the gear and pinion are of equal diameter, of identical shape, which is an obvious advantage from the manufacturing standpoint.

# LETTERS TO THE EDITOR

*This section will accept brief reports of new work of general interest to industrial physicists. The Editorial Board will not hold itself responsible for opinions expressed by the correspondents. Contributions to this section must reach the office of the Managing Editor not later than the twenty-fifth of the month preceding that of the issue in which the letter is to appear.*

## A Formula for the Birefringence of Vibrating Media

Consider an optical arrangement (such as used by Günther[1]) in which a transparent medium is placed between a pair of crossed Nicols and is traversed by a beam of parallel monochromatic light. The diagonals of the Nicols are set at 45° with respect to the principal plane of the medium. When the medium is caused to vibrate, the field of vision brightens and/or darkens in those regions in which the mechanical strain is sufficiently great. The following theory, which has been used by the writers for the past two years, permits a quantitative analysis of the vibrations of transparent media, as for example piezoelectric crystals.

The intensity $I$ of the light transmitted through both Nicols is[2]

$$I = I_{\max} \sin^2 \pi R_\lambda, \tag{1}$$

where $R_\lambda$ is the retardation of the extraordinary ray with respect to the ordinary ray. If the medium be subjected to a force field $F \sin (2\pi t/T)$, one can show from Pockels' theories[3] for the *linear* piezo- and electro-optical effects that

$$R_\lambda = cLF \sin (2\pi t/T), \tag{2}$$

where $L$ is the thickness of the medium and $c$ is a linear function of the optical constants. Combining Eqs. (1) and (2) one now finds the *instantaneous* intensity to be

$$\begin{aligned} I_i &= I_{\max} \sin^2 (\pi cLF \sin (2\pi t/T)) \\ &= (I_{\max}/2)[1 - \cos (2\pi cLF \sin 2\pi t/T)] \\ &\equiv (I_{\max}/2)[1 - \cos (x \sin \phi)]. \end{aligned} \tag{3}$$

When the frequency $1/T$ is more rapid than 20 cycles per second, the eye discerns the average intensity $I_a$.

$$\begin{aligned} I_a &= \frac{I_{\max}}{2} \frac{2}{T} \int_0^{T/2} [1 - \cos (x \sin \phi)] dt \\ &= \frac{I_{\max}}{2\pi} \int_0^{\pi} [1 - \cos (x \sin \phi)] d\phi \\ &= \frac{I_{\max}}{2} [1 - J_0(x)]^4 = \frac{I_{\max}}{2} [1 - J_0(2\pi cLF)]. \end{aligned} \tag{4}$$

As $F$, and therefore $x$, is increased, $I_a$ increases until it reaches a maximum at $x \doteq 3.83$. With increasing values of $x$, $I_a$ decreases to a minimum at $x \doteq 7.02$, reaches another maximum at $x \doteq 10.17$, a second minimum at $x \doteq 13.32$, etc. For every observed maximum or minimum the corresponding value of $x$ is thus known.

This theory explains the beautiful photographs by Brüninghaus.[5] It also explains his "polarization chromatique" when one includes the variation of $c$ with wavelength.

The first minimum of $I$, Eq. (1), is reached in the *static case* when $R_\lambda = 1$ or $F = 1/cL$. The corresponding minimum is reached in the *dynamic case* when $F = (7.02/2\pi)(1/cL) = 1.12/cL$, or at a value 12 percent greater than in the static case.

For values of $x$ less than 1, $I_a$ is given with an accuracy greater than 93 percent by

$$I_a = I_{\max} x^2/8.$$

This equation predicts the intensity of light which passes through the Okolicsanyi cell,[6, 7] an instrument used in television.

Papers relating to the application of this method to special problems in photoelasticity will be published elsewhere.

JOHN W. COOKSON
HAROLD OSTERBERG

University of Wisconsin,
Madison, Wisconsin,
March 25, 1936.

[1] N. Günther, Ann. d. Physik **405**, 792 (1932).
[2] Houston, *A Treatise on Light* (London, 1927), p. 199
[3] F. Pockels, *Lehrbuch der Kristalloptik* (Leipzig, 1906), Part IV.
[4] Gray, Matthews and MacRoberts, *Bessel Functions* (London, 1922), p. 32.
[5] Brüninghaus, J. de physique **6**, 159 (1935).
[6] F. v. Okolicsanyi, Zeits. Fernsehen u. Tonfilm, Aug. (1933), (Berlin 4. Jahrg. No. 4, 43).
[7] Ch. Schramm, Ann. d. Physik **25**, 309 (1936).